

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2004, 06:44:49 ; Search time 2545 Seconds  
(without alignments)  
9504.270 Million cell updates/sec

Title: US-09-852-100B-1  
Perfect score: 810  
Sequence: 1 atgcatatttttaaagggtc.....aaacgcaattatatccataa 810

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estmu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_htc:\*
- 9: gb\_est1:\*
- 10: gb\_est2:\*
- 11: gb\_htc:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_est5:\*
- 15: em\_estfun:\*
- 16: em\_estom:\*
- 17: em\_gss\_hum:\*
- 18: em\_gss\_inv:\*
- 19: em\_gss\_pln:\*
- 20: em\_gss\_vrt:\*
- 21: em\_gss\_fun:\*
- 22: em\_gss\_mam:\*
- 23: em\_gss\_mus:\*
- 24: em\_gss\_pro:\*
- 25: em\_gss\_rod:\*
- 26: em\_gss\_phg:\*
- 27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	%		Query			ID	Description
	Score	Match	Length	DB			
1	673	83.1	678	12	BG702759		BG702759 602684629
2	643.8	79.5	836	12	BG723403		BG723403 602694073
3	642.8	79.4	788	14	CB996712		CB996712 AGENCOURT
4	636.4	78.6	658	12	BG709182		BG709182 602675061
5	635	78.4	982	11	BC048995		BC048995 Homo sapi
6	631.6	78.0	750	12	BI458114		BI458114 603198535
7	626.2	77.3	975	12	BI546941		BI546941 603190155
8	622.4	76.8	882	12	BI464436		BI464436 603205310
9	621.6	76.7	879	12	BI462204		BI462204 603205517
10	613.4	75.7	615	13	BQ639765		BQ639765 he20a04.y
11	606.4	74.9	772	14	CB310671		CB310671 AGENCOURT
12	606	74.8	950	12	BI562596		BI562596 603256530
13	605.4	74.7	901	12	BI596830		BI596830 603243323
14	604.4	74.6	908	12	BI596662		BI596662 603243232
15	595.4	73.5	599	9	AI923178		AI923178 wn67b10.x
16	587.4	72.5	788	14	CB229262		CB229262 AGENCOURT
17	576.8	71.2	775	14	CB311691		CB311691 AGENCOURT
18	572.4	70.7	574	9	AI680904		AI680904 tx42f05.x
19	571	70.5	961	10	BF968960		BF968960 602270156
20	556	68.6	943	13	BQ232033		BQ232033 AGENCOURT
21	555	68.5	556	9	AI479764		AI479764 tm69b04.x
22	547	67.5	547	10	BE467256		BE467256 hz63f03.x
23	534	65.9	534	10	BE467235		BE467235 hz63d03.x
24	531.2	65.6	537	9	AI580361		AI580361 tm47h02.x
25	530.8	65.5	534	10	BE328144		BE328144 ht97f11.x
26	527.4	65.1	530	9	AI674462		AI674462 wc44e01.x
27	521.4	64.4	523	12	BM698864		BM698864 UI-E-DX1-
28	516	63.7	516	10	BE348755		BE348755 ht70b01.x
29	513.8	63.4	542	12	BG700625		BG700625 602682226
30	513.4	63.4	548	9	AA772225		AA772225 ai41c01.s
31	513.4	63.4	771	13	BX108300		BX108300 BX108300
32	506	62.5	507	9	AI682204		AI682204 wa71b06.x
33	503.6	62.2	603	9	AV654663		AV654663 AV654663
34	497	61.4	498	9	AI349520		AI349520 qp72g05.x
35	495.6	61.2	526	10	BE549799		BE549799 7b38g08..x
36	494.4	61.0	506	14	CB122011		CB122011 K-EST0169
37	492.8	60.8	500	9	AI680969		AI680969 tx35h07.x
38	476	58.8	870	14	CB993944		CB993944 AGENCOURT
39	475.2	58.7	799	12	BG701162		BG701162 602680739
40	474.6	58.6	586	14	CB430834		CB430834 606754 MA
41	472.8	58.4	808	14	CB991349		CB991349 AGENCOURT
42	471.2	58.2	830	14	CB992200		CB992200 AGENCOURT
43	466.2	57.6	791	14	CA451757		CA451757 UI-M-FX0-
44	464	57.3	891	10	BF692101		BF692101 602247882
45	459.2	56.7	922	13	BQ876648		BQ876648 AGENCOURT

## ALIGNMENTS

## RESULT 1

BG702759

LOCUS BG702759 678 bp mRNA linear EST 07-MAY-2001

DEFINITION 602684629F1 NIH\_MGC\_95 Homo sapiens cDNA clone IMAGE:4817358 5', mRNA sequence.

ACCESSION BG702759

VERSION BG702759.1 GI:13974418

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 678)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki

Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM10717 row: i column: 07

High quality sequence stop: 678.

## FEATURES

source

Location/Qualifiers

1. .678

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:4817358"

/tissue\_type="hippocampus"

/lab\_host="DH10B"

/clone\_lib="NIH\_MGC\_95"

/note="Organ: brain; Vector: pBluescriptR (modified

pBluescript KS+); Site\_1: BamHI; Site\_2: SalI-XhoI

(gtcgag); Oligo-dT primed using primer

5'-TTTTTTTTTTTTTTTTVN-3', size-selected for average

insert size 2.5 kb and normalized to ROT 5. This is a

primary library enriched for full-length clones and

constructed using the Cap-trapper method (Carninci, in

preparation). Library constructed by M. Brownstein

(NIMH/NHGRI, National Institutes of Health). Note: this

is a NIH\_MGC Library."

## ORIGIN

Query Match 83.1%; Score 673; DB 12; Length 678;

Best Local Similarity 100.0%; Pred. No. 4.7e-175;

Matches 673; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	135	GCTCCTGGGCGGAGGCGGAAGCGGAAGTGGCGAGAAAGTGTCTGGTCTCCAAGATGGCGGC	194
Db	6	GCTCCTGGGCGGAGGCGGAAGCGGAAGTGGCGAGAAAGTGTCTGGTCTCCAAGATGGCGGC	65
Qy	195	CGCCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGGCCAGACTCGTTGGTGTCT	254
Db	66	CGCCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGGCCAGACTCGTTGGTGTCT	125
Qy	255	GTGGTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGCCACCTCCGCCGGGGGCGA	314
Db	126	GTGGTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGCCACCTCCGCCGGGGGCGA	185
Qy	315	GGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAAAAT	374
Db	186	GGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAAAAT	245
Qy	375	AAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCTCATGTTTCCTGTTT	434
Db	246	AAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCTCATGTTTCCTGTTT	305
Qy	435	TCCAGCACCCAACATAACTTGTAAAGGATTCAGTGGCAATGAAACACATTTTACTGGGAA	494
Db	306	TCCAGCACCCAACATAACTTGTAAAGGATTCAGTGGCAATGAAACACATTTTACTGGGAA	365
Qy	495	CGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCCTACAAAGT	554
Db	366	CGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCCTACAAAGT	425
Qy	555	GGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGGATA	614
Db	426	GGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGGATA	485
Qy	615	CCCTGCTTTGGGTTTGTAAAGTTTGCAGTGTAGGGTTTGTGGAATTGGGAGCCTAAT	674
Db	486	CCCTGCTTTGGGTTTGTAAAGTTTGCAGTGTAGGGTTTGTGGAATTGGGAGCCTAAT	545
Qy	675	TGATTTCACTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACATTAT	734
Db	546	TGATTTCACTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACATTAT	605
Qy	735	AGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAAAAC	794
Db	606	AGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAAAAC	665
Qy	795	GCAATTATATCCA	807
Db	666	GCAATTATATCCA	678

# RESULT 2

BG723403

LOCUS

BG723403

836 bp

mRNA

linear

EST 08-MAY-2001

DEFINITION

602694073F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:4826035 5', mRNA sequence.

ACCESSION

BG723403

VERSION

BG723403.1 GI:14002590

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 836)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
Toshiyuki and Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM10740 row: b column: 20  
High quality sequence stop: 760.

FEATURES Location/Qualifiers

source 1. .836  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4826035"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_97"  
/note="Organ: testis; Vector: pBluescriptR (modified  
pBluescript KS+); Site\_1: BamHI; Site\_2: SalI-XhoI  
(gtcgag); Oligo-dT primed using primer  
5'-TTTTTTTTTTTTTTTVN-3', size-selected for average  
insert size 2.2 kb and normalized to ROT 5. This is a  
primary library enriched for full-length clones and  
constructed using the Cap-trapper method (Carninci, in  
preparation). Library constructed by M. Brownstein  
(NIMH/NHGRI, National Institutes of Health). Note: this is  
a NIH\_MGC Library."

# ORIGIN

Query Match 79.5%; Score 643.8; DB 12; Length 836;  
Best Local Similarity 99.7%; Pred. No. 6.2e-167;  
Matches 645; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	164	GCGAGAAAGTGTCTGGTCTCCAAGATGGCGGCCCGCTGGCCGTCTGGTCCGTCTGCTCCGG	223
Db	2	GCGGAAAAGTGTCTGGTCTCCAAGATGGCGGCCCGCTGGCCGTCTGGTCCGTCTGCTCCGG	61
Qy	224	AGGCCGTGACGGCCAGACTCGTTGGTGTCTGTGGTTCGTCTCAGTCACTACAGGACCCT	283
Db	62	AGGCCGTGACGGCCAGACTCGTTGGTGTCTGTGGTTCGTCTCAGTCACTACAGGACCCT	121
Qy	284	GGGGGGCTGTTGCCACCTCCGCCGGGGGCGAGGAGTCGCTTAAGTGCAGGACCTCAAAG	343
Db	122	GGGGGGCTGTTGCCACCTCCGCCGGGGGCGAGGAGTCGCTTAAGTGCAGGACCTCAAAG	181

Qy 344 TGGGACAATATATTTGTAAAGATCCAAAAATAAATGACGCTACGCAAGAACCAGTTAACT 403  
 |||  
 Db 182 TGGGACAATATATTTGTAAAGATCCAAAAATAAATGACGCTACGCAAGAACCAGTTAACT 241  
 Qy 404 GTACAAACTACACAGCTCATGTTTCCTGTTTTCCAGCACCCAACATAACTTGTAAGGATT 463  
 |||  
 Db 242 GTACAAACTACACAGCTCATGTTTCCTGTTTTCCAGCACCCAACATAACTTGTAAGGATT 301  
 Qy 464 CCAGTGGCAATGAAACACATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTT 523  
 |||  
 Db 302 CCAGTGGCAATGAAACACATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTT 361  
 Qy 524 GCCGAAATGTAAATGGCTATTTCCTACAAAGTGGCAGTCGCATTGTCTCTTTTTCTTGGAT 583  
 |||  
 Db 362 GCCGAAATGTAAATGGCTATTTCCTACAAAGTGGCAGTCGCATTGTCTCTTTTTCTTGGAT 421  
 Qy 584 GGTGGGAGCAGATCGATTTTACCTTGGATACCCTGCTTTGGGTTTGTTAAAGTTTTCGA 643  
 |||  
 Db 422 GGTGGGAGCAGATCGATTTTACCTTGGATACCCTGCTTTGGGTTTGTTAAAGTTTTCGA 481  
 Qy 644 CTGTAGGGTTTTGTGGAATTGGGAGCCTAATTGATTTTCATTCTTATTTCAATGCAGATTG 703  
 |||  
 Db 482 CTGTAGGGTTTTGTGGAATTGGGAGCCTAATTGATTTTCATTCTTATTTCAATGCAGATTG 541  
 Qy 704 TTGGACCTTCAGATGGAAGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGAC 763  
 |||  
 Db 542 TTGGACCTTCAGATGGAAGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGAC 601  
 Qy 764 TGAGTATTACTAATGAAACATTTAGAAAAACGCAATTATATCCATAA 810  
 |||  
 Db 602 TGAGTATTACTAATGAAACATTTAGAAAAACGCAATTATATCCATAA 648

RESULT 3

CB996712

LOCUS CB996712 788 bp mRNA linear EST 01-MAY-2003

DEFINITION AGENCOURT\_13627955 NIH\_MGC\_148 Homo sapiens cDNA clone  
 IMAGE:30334410 5', mRNA sequence.

ACCESSION CB996712

VERSION CB996712.1 GI:30291232

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 788)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: Dr. Stefan Hansson

cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help  
 and advice from Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

```

FEATURES
    source
        Location/Qualifiers
            1. .788
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:30334410"
                /tissue_type="pre-eclamptic placenta"
                /lab_host="DH10B TonA"
                /clone_lib="NIH_MGC_148"
                /note="Organ: placenta; Vector: pBluescriptR; Site_1:
                all-XhoI; Site_2: BamH; Library is oligo-dT primed and
                directionally cloned using primer
                5'-TTTTTTTTTTTTTTTTTVN-3', size-selected for average insert
                size 2.3 kb and normalized to ROT 5. This is a primary
                library enriched for full-lenght clones and constructed
                using the Cap-trapper method (Carninci, in preparation).
                Library constructed by M. Brownstein (NIMH/NHGRI,
                National Institutes of Health). Note: this is a NIH_MGC
                Library."

```

Query Match 79.4%; Score 642.8; DB 14; Length 788;  
Best Local Similarity 99.7%; Pred. No. 1.1e-166;  
Matches 644; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

[illegible]

Qy 585 GTTGGGAGCAGATCGATTTTACCTTGGATACCCTGCTTGGGTTTGTAAAGTTTGCAC 644  
 |||  
 Db 444 GTTGGGAGCAGATCGATTTTACCTTGGATACCCTGCTTGGGTTTGTAAAGTTTGCAC 503  
 |||  
 Qy 645 TGTAGGGTTTTGTGGAATTGGGAGCCTAATTGATTTCATTCTTATTTCAATGCAGATTGT 704  
 |||  
 Db 504 TGTAGGGTTTTGTGGAATTGGGAGCCTAATTGATTTCATTCTTATTTCAATGCAGATTGT 563  
 |||  
 Qy 705 TGGACCTTCAGATGGAAGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACT 764  
 |||  
 Db 564 TGGACCTTCAGATGGAAGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACT 623  
 |||  
 Qy 765 GAGTATTACTAATGAAACATTTAGAAAAACGCAATTATATCCATAA 810  
 |||  
 Db 624 GAGTATTACTAATGAAACATTTAGAAAAACGCAATTGTATCCATAA 669  
 |||

# RESULT 4

BG709182

LOCUS BG709182 658 bp mRNA linear EST 07-MAY-2001

DEFINITION 602675061F1 NIH\_MGC\_96 Homo sapiens cDNA clone IMAGE:4797782 5', mRNA sequence.

ACCESSION BG709182

VERSION BG709182.1 GI:13987263

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 658)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM10684 row: i column: 15

High quality sequence stop: 658.

## FEATURES

Location/Qualifiers

source

1..658

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:4797782"

/tissue\_type="hypothalamus"

/lab\_host="DH10B"

/clone\_lib="NIH\_MGC\_96"

/note="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site\_1: BamHI; Site\_2: SalI-XhoI (gtcgag); Oligo-dT primed using primer





## RESULT 5

BC048995

LOCUS BC048995 982 bp mRNA linear HTC 17-DEC-2003

DEFINITION Homo sapiens beta-amyloid binding protein precursor, mRNA (cDNA clone IMAGE:5261702).

ACCESSION BC048995

VERSION BC048995.1 GI:28981340

KEYWORDS HTC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 982)

AUTHORS

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

PUBMED 12477932

REFERENCE 2 (bases 1 to 982)

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (17-MAR-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)~~Tissue Procurement: Miklos Palkovits, M.D., Ph.D.~~

cDNA Library Preparation: Michael J. Brownstein (NHGRI) &amp; Shiraki Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Institute for Systems Biology

<http://www.systemsbiology.org>contact: [amadan@systemsbiology.org](mailto:amadan@systemsbiology.org)

Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found



Qy 708 ACCTTCAGATGGAAGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAG 767  
 |||  
 Db 541 ACCTTCAGATGGAAGTAGTTACATTATAGATTACTAAGGAACCAGACTTACAAGACTGAG 600  
 Qy 768 TATTACTAATGAAACATTTAGAAAAACGCAATTATATCCATAA 810  
 |||  
 Db 601 TATTACTAATGAAACATTTAGAAAAACGCAATTATATCCATAA 643

RESULT 6

BI458114

LOCUS BI458114 750 bp mRNA linear EST 21-AUG-2001

DEFINITION 603198535F1 NIH\_MGC\_96 Homo sapiens cDNA clone IMAGE:5278064 5',  
 mRNA sequence.

ACCESSION BI458114

VERSION BI458114.1 GI:15248770

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 750)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
 Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM11702 row: e column: 09

High quality sequence stop: 711.

FEATURES

source

Location/Qualifiers

1. .750

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5278064"

/tissue\_type="hypothalamus"

/lab\_host="DH10B"

/clone\_lib="NIH\_MGC\_96"

/note="Organ: brain; Vector: pBluescriptR (modified

pBluescript KS+); Site\_1: BamHI; Site\_2: SalI-XhoI

(gtcgag); Oligo-dT primed using primer

5'-TTTTTTTTTTTTTTTVN-3', size-selected for average

insert size 2.3 kb and normalized to ROT 5. This is a

primary library enriched for full-length clones and

constructed using the Cap-trapper method (Carninci, in

preparation). Library constructed by M. Brownstein

(NIMH/NHGRI, National Institutes of Health). Note: this is

a NIH\_MGC Library."

ORIGIN





Qy 340 AAAGTGGGACAATATATTTGTAAAGATCCAAAAATAAATGACGCTACGCAAGAACCAGTT 399  
 |||  
 Db 181 AAAGTGGGACAATATATTTGTAAAGATCCAAAAATAAATGACGCTACGCAAGAACCAGTT 240  
 |||  
 Qy 400 AACTGTACAACTACACAGCTCATGTTTCCTGTTTTCCAGCACCCAACATAACTTGTAAG 459  
 |||  
 Db 241 AACTGTACAACTATACAGCTCATGTTTCCTGTTTTCCAGCACACAACATAACTTGTAAG 300  
 |||  
 Qy 460 GATTCCAGTGGCAATGAAACACATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATA 519  
 |||  
 Db 301 GATTCCAGTGGCAATGAAACACATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATA 360  
 |||  
 Qy 520 TCTTGCCGAAATGTAAATGGCTATTCTACAAAGTGGCAGTCGCATTGTCTCTTTTCTT 579  
 |||  
 Db 361 TCTTGCCGAAATGTAAATGGCTATTCTACAAAGTGGCAGTAGCATTGTCTCTTTTCTT 420  
 |||  
 Qy 580 GGATGGTTGGGAGCAGATCGATTTTACCTTGGATACCCTGCTTTGGGTTTGTAAAGTTT 639  
 |||  
 Db 421 GGATGGTTGGGAGCAGATCGATTTTACCTTGGATACCCTGCTTTGGGTTTGTAAAGTTT 480  
 |||  
 Qy 640 TGCAGTGTAGGGTTTTGTGGAATTGGGAGCCTAATTGATTTTATTCTTATTTCATGCAG 699  
 |||  
 Db 481 TGCAGTGTAGGGTTTTGTGGAATTGGGAGCCTAATTGATTTTATTCTTATTTCATGCAG 540  
 |||  
 Qy 700 ATTGTTGGACCTTCAGATGGAAGTAGTTACATTATAGATTACTATGGAACCAGACTTACA 759  
 |||  
 Db 541 ATTGTTGGACCTTCAGATGGAAGTAGTTACATTATAGATTACTAAGGAACCAGACTTACA 600  
 |||  
 Qy 760 AGACTGAGTATTACTAATGAAACATTTAGAAAAACGCAATTATATCCATAA 810  
 |||  
 Db 601 AGACTGAGTATTACTAATGAAACA-TTAGAAAAACGCAATTATATCCATAA 650  
 |||

RESULT 8

BI464436

LOCUS BI464436 882 bp mRNA linear EST 21-AUG-2001  
 DEFINITION 603205310F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:5271098 5',  
 mRNA sequence.

ACCESSION BI464436

VERSION BI464436.1 GI:15255092

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 882)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
 Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.





Qy 584 GGTGGGAGCAGATCGATTTTACCTTGGATACCCTGCTTTGGGTTTGTAAAGTTTGGCA 643  
 |||  
 Db 424 GGTGGGAGCAGATCGATTTTACCTTGGATACCCTGCTTTGGGTTTGTAAAGTTTGGCA 483  
 Qy 644 CTGTAGGGTTTTGTGGAATTGGGAGCCTAATTGATTTTCATTCTTATTTCAATGCAGATTG 703  
 |||  
 Db 484 CTGTAGGG-TTTGTGGAATTGGGAGCCTAATTGATTTTCATTCTTATTTCAATGCAGATTG 542  
 Qy 704 TTGGACCTTCAGATGGAAGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGAC 763  
 |||  
 Db 543 TTGGACCTTCAGATGGAAGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGAC 602  
 Qy 764 TGAGTATTACTAATGAAACA-TTTAGAAAAACGCAATTATATCCATAA 810  
 |||  
 Db 603 TGAGTATTACTAATGAAACATTTTAGAAAAACGCAATTATATCCATAA 650

RESULT 9

BI462204

LOCUS BI462204 879 bp mRNA linear EST 21-AUG-2001

DEFINITION 603205517F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:5271077 5', mRNA sequence.

ACCESSION BI462204

VERSION BI462204.1 GI:15252860

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 879)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM11684 row: b column: 06

High quality sequence stop: 753.

FEATURES

source

Location/Qualifiers

1..879

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5271077"

/lab\_host="DH10B"

/clone\_lib="NIH MGC\_97"

/note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site\_1: BamHI; Site\_2: SalI-XhoI

(gtcgag); Oligo-dT primed using primer

5'-TTTTTTTTTTTTTTTTVN-3', size-selected for average



## RESULT 10

BQ639765

LOCUS BQ639765 615 bp mRNA linear EST 15-JUL-2002

DEFINITION he20a04.y1 Human Retina cDNA (Un-normalized, unamplified): hd/he  
Homo sapiens cDNA clone he20a04 5', mRNA sequence.

ACCESSION BQ639765

VERSION BQ639765.1 GI:21764224

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 615)

AUTHORS Wistow,G., Bernstein,S.L., Wyatt,M.K., Ray,S., Behal,A.,

Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.

TITLE Expressed sequence tag analysis of human retina for the NEIBank  
Project: Retbindin, an abundant, novel retinal cDNA and alternative  
splicing of other retina-preferred gene transcripts

JOURNAL Mol. Vis. 8 (4), 196-204 (2002)

MEDLINE 22103461

PUBMED 12107411

COMMENT Contact: Wistow G

Section on Molecular Structure and Function

National Eye Institute

6/331, NIH, Bethesda, MD 20892-2740, USA

Tel: 301 402 3452

Fax: 301 496 0078

Email: graeme@helix.nih.gov

Plate: 20 row: a column: 04

Seq primer: M13RP1 reverse primer (ABI).

## FEATURES

source

Location/Qualifiers

1. .615

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="he20a04"

/tissue\_type="Retina"

/dev\_stage="Adult"

/lab\_host="EMDH10B"

/clone\_lib="Human Retina cDNA (Un-normalized,  
unamplified): hd/he"/note="Organ: Eye; Vector: pSPORT1; Neural retina tissue  
was dissected from two 80 year old donors with no observed  
eye disease. 100ug of total RNA was used for library  
construction. A directionally cloned cDNA library in the  
pSPORT1 vector (Life Technologies) was constructed at  
Bioserve Biotechnology (Laurel MD) essentially following  
the protocols of the SuperScript Plasmid System full  
details of which are contained in the manufacturer's  
Instruction manual (<http://www.lifetech.com/>). First  
strand synthesis was carried out using a Not I  
primer-adapter[5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. EST analysis  
was performed on the unamplified library at the NIH  
Intramural Sequencing Center (NISC)."

ORIGIN

Query Match 75.7%; Score 613.4; DB 13; Length 615;  
 Best Local Similarity 99.8%; Pred. No. 1.5e-158;  
 Matches 614; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy      193 GCCGCCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGCCAGACTCGTTGGTGTC 252
          |||
Db      1   GCCGCCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGCCAGACTCGTTGGTGTC 60

Qy      253 CTGTGGTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGCCACCTCCGCCGGGGGC 312
          |||
Db      61   CTGTGGTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGCCACCTCCGCCGGGGGC 120

Qy      313 GAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAAA 372
          |||
Db      121 GAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAAA 180

Qy      373 ATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCTCATGTTTCCTGT 432
          |||
Db      181 ATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCTCATGTTTCCTGT 240

Qy      433 TTTCCAGCACCCAACATAAATTGTAAGGATTCAGTGGCAATGAAACACATTTTACTGGG 492
          |||
Db      241 TTTCCAGCACCCAACATAAATTGTAAGGATTCAGTGGCAATGAAACACATTTTACTGGG 300

Qy      493 AACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCCTACAAA 552
          |||
Db      301 AACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCCTACAAA 360

Qy      553 GTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGA 612
          |||
Db      361 GTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGA 420

Qy      613 TACCCTGCTTTGGGTTTGTAAAGTTTTGCACTGTAGGGTTTTGTGGAATTGGGAGCCTA 672
          |||
Db      421 TACCCTGCTTTGGGTTTGTAAAGTTTTGCACTGTAGGGTTTTGTGGAATTGGGAGCCTA 480

Qy      673 ATTGATTTCACTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACATT 732
          |||
Db      481 ATTGATTTCACTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACATT 540

Qy      733 ATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAAA 792
          |||
Db      541 ATAGATTACTATGGGACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAAA 600

Qy      793 ACGCAATTATATCCA 807
          |||
Db      601 ACGCAATTATATCCA 615
  
```

RESULT 11

CB310671

LOCUS CB310671 772 bp mRNA linear EST 04-MAR-2003

DEFINITION AGENCOURT\_11828318 NICHD\_Rh\_Ov1 Macaca mulatta cDNA clone

IMAGE:6895132 5', mRNA sequence.

ACCESSION CB310671

VERSION CB310671.1 GI:28833385

KEYWORDS EST.



Qy 403 TGTACAAACTACACAGCTCATGTTTCCTGTTTCCAGCACCCAACATAACTTGTAAGGAT 462  
 |||  
 Db 241 TGTACAAACTACACAGCTCATGTTTCCTGTTTCCAGCACCTAACATAACTTGTAAGGAT 300  
 Qy 463 TCCAGTGGCAATGAAACACATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCT 522  
 |||  
 Db 301 TCCAGTGGCAATGAAACACATTTTACTGGGAATGAAGTTGGTTTTTTCAAGCCCATATCT 360  
 Qy 523 TGCCGAAATGTAAATGGCTATTCTACAAAGTGGCAGTCGCATTGTCTCTTTTCTTGGA 582  
 |||  
 Db 361 TGCCGAAATGTAAATGGCTATTCTACAAAGTGGCAGTCGCATTGTCTCTTTTCTTGGA 420  
 Qy 583 TGGTTGGGAGCAGATCGATTTTACCTTGGGATACCCTGCTTTGGGTTTGTAAAGTTTGC 642  
 |||  
 Db 421 TGGTTGGGAGCAGATCGATTTTACCTGGGATACCCTGCCTTGGGTTTGTAAAGTTTGC 480  
 Qy 643 ACTGTAGGGTTTTGTGGAATTGGGAGCCTAATTGATTTTATTCTTATTTCATGCAGATT 702  
 |||  
 Db 481 ACTGTAGGATTTTGTGGAATTGGGAGCCTAATTGATTTTATTCTTATTTCATGCAGATT 540  
 Qy 703 GTTGGACCTTCAGATGGAAGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGA 762  
 |||  
 Db 541 GTTGGACCTTCAGATGGAAGTAGTTACATCATAGATTACTATGGAACCAGACTGACAAGA 600  
 Qy 763 CTGAGTATTACTAATGAAACATTTAGAAAAACGCAATTATATCCATAA 810  
 ||  
 Db 601 CTAAGTATTACTAATGAAACATATAGAAAAACGCAATTATATCCATAA 648

# RESULT 12

BI562596

LOCUS BI562596 950 bp mRNA linear EST 05-SEP-2001

DEFINITION 603256530F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:5298943 5', mRNA sequence.

ACCESSION BI562596

VERSION BI562596.1 GI:15449910

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 950)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM11756 row: k column: 08

High quality sequence stop: 753.



```

      |||
Db      483 GCACTGTAGGG-TTTGTGGAATTGGGAGCCTAATTGATTCATTCTTATTTCAATGCAAG 541
Qy      699 GATTGTTGGACCTTCAGATGGAAGTAGTTACATTATAGATTACTATGGAACCAGACTTAC 758
      | |||
Db      542 ATTGGTTGGACCTTCAGATGGAAGTAGTTACATTATAGATTACTATGGAACCAGACTTAC 601
Qy      759 AAGACTGAGTATTACTAATGAAACATTTAGAAAAACGCAATTATATCCATAA 810
      |||
Db      602 AAGACTGAGTATTACTAATGAAACATTTAGAAAAACGCAATTATATCCATAA 653

```

# RESULT 13

BI596830

LOCUS BI596830 901 bp mRNA linear EST 07-SEP-2001

DEFINITION 603243323F1 NIH\_MGC\_96 Homo sapiens cDNA clone IMAGE:5285933 5', mRNA sequence.

ACCESSION BI596830

VERSION BI596830.1 GI:15489769

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 901)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM11722 row: m column: 06

High quality sequence stop: 760.

## FEATURES

source

Location/Qualifiers

1. .901

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5285933"

/tissue\_type="hypothalamus"

/lab\_host="DH10B"

/clone\_lib="NIH\_MGC\_96"

/note="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site\_1: BamHI; Site\_2: Sali-XhoI (gtcgag); Oligo-dT primed using primer

5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein







Qy	292	GTTGCCACCTCCGCCGGGGGCGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAA	351
Db	124	GTTGCCACCTCCGCCGGGGGCGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAA	183
Qy	352	TATATTTGTAAAGATCCAAAAATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAAC	411
Db	184	TATATTTGTAAAGATCCAAAAATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAAC	243
Qy	412	TACACAGCTCATGTTTCCTGTTTTCCAGCACCCAACATAACTT-GTAAGGATTCCAGTGG	470
Db	244	TACACAGCTCATGTTTCCTGTTTTCCAGCACCCAACATAACTTGGTAAGGATTCCAGTGG	303
Qy	471	CAATGAAACACATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAA	530
Db	304	CAATGAAACACATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAA	363
Qy	531	TGTAAATGGCTATTCTACAAAGTGGCAGTCGCATTGTCTCTTTTTCTTGATGGTTGGG	590
Db	364	TGTAAATGGCTATTCTACAAAGTGGCAGTCGCATTGTCTCTTTTTCTTGATGGTTGGG	423
Qy	591	AGCAGATCGATTTTACCTTGGATACCCTGCTTTGGGTTTGTAAAGTTTGCAGTGTAGG	650
Db	424	AGCAGATCGATTTTACCTTGGATACCCTGCTTTGGGTTTGTAAAGTTTGCAGTGTAGG	483
Qy	651	GTTTTGTGGAATTGGGAGCCTAATTGATTTTCATTCTTATTTCAATGCAGATTGTTGGACC	710
Db	484	GTTT--GTGGAATGGGAGCCTAATTGATTTTCATTCTTATTTCAATGCAGATTGTTGGACC	541
Qy	711	TTCAGATGGAAGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTAT	770
Db	542	TTCAGATGGAAGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTAT	601
Qy	771	TACTAATGAAACATTTAGAAAAACGCAATTATATCCATAA	810
Db	602	TACTAATGAAACATTTAGAAAAACGCAATTATATCCATAA	641

# RESULT 15

AI923178

LOCUS AI923178 599 bp mRNA linear EST 02-SEP-1999  
 DEFINITION wn67b10.x1 NCI CGAP Lu19 Homo sapiens cDNA clone IMAGE:2450491 3' similar to WP:C02F5.3 CE00039 GTP-BINDING PROTEIN ;, mRNA sequence.

ACCESSION AI923178

VERSION AI923178.1 GI:5659142

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 599)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cdNA Library Preparation: M. Bento Soares, Ph.D.

cdNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)

Seq primer: -40UP from Gibco

High quality sequence stop: 457.

#### FEATURES

source

Location/Qualifiers

1. .599

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:2450491"

/tissue\_type="squamous cell carcinoma, poorly differentiated (4 pooled tumors, including primary and metastatic)"

/dev\_stage="adult"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NCI\_CGAP\_Lu19"

/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from pooled lung tumor tissue, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo. "

#### ORIGIN

Query Match 73.5%; Score 595.4; DB 9; Length 599;

Best Local Similarity 99.5%; Pred. No. 1.4e-153;

Matches 596; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 190 GCGGCCGCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGGCCAGACTCGTTGGT 249  
|||||

Db 1 GCGGCCGCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGGCCAGACTCGTTGGT 60

Qy 250 GTCCTGTGGTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGCCACCTCCGCCGGG 309  
|||||

Db 61 GTCCTGTGGTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGCCACCTCCGCCGGG 120

Qy 310 GGCGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCA 369  
|||||

Db 121 GGCGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCA 180

Qy 370 AAAATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCTCATGTTTCC 429  
|||||

Db 181 AAAATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCTCATGTTTCC 240

Qy 430 TGTTTTCCAGCACCCACATAACTTGTAAGGATTCCAGTGGCAATGAAACACATTTTACT 489  
|||||

Db 241 TGTTTTCCAGCACCCACATAACTTGTAAGGATTCCAGTGGCAATGAAACACATTTTACT 300

Qy 490 GGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCCTAC 549

```

      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 GGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCCTAC 360
Qy      550 AAAGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTT 609
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 AAAGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTT 420
Qy      610 GGATACCCTGCTTTGGGTTTGTAAAGTTTTGCACTGTAGGGTTTTGTGGAATTGGGAGC 669
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 GGATACCCTGCTTTGGGTTTGTAAAGTTTTGCACTGTANGGTTTTGTGGAATTGGGAGC 480
Qy      670 CTAATTGATTTCAATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTAC 729
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 CTAATTGATTTCAATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTAC 540
Qy      730 ATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAG 788
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 ATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATNTAG 599

```

Search completed: March 4, 2004, 09:16:38  
Job time : 2551 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2004, 05:35:01 ; Search time 3285 Seconds  
(without alignments)  
10687.323 Million cell updates/sec

Title: US-09-852-100B-1  
Perfect score: 810  
Sequence: 1 atgcatatttttaaaagggtc.....aaacgcaattatatccataa 810

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*

28: em\_un:\*  
 29: em\_vi:\*  
 30: em\_htg\_hum:\*  
 31: em\_htg\_inv:\*  
 32: em\_htg\_other:\*  
 33: em\_htg\_mus:\*  
 34: em\_htg\_pln:\*  
 35: em\_htg\_rod:\*  
 36: em\_htg\_mam:\*  
 37: em\_htg\_vrt:\*  
 38: em\_sy:\*  
 39: em\_htgo\_hum:\*  
 40: em\_htgo\_mus:\*  
 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

8									
Result	Query								
No.	Score	Match	Length	DB	ID	Description			
1	810	100.0	810	6	BD243134	BD243134	6-Protein		
2	810	100.0	1246	9	AF353990	AF353990	Homo sapi		
3	645	79.6	984	9	BC029486	BC029486	Homo sapi		
4	602.8	74.4	970	6	BD139411	BD139411	Extended		
5	499	61.6	508	6	BD265227	BD265227	Compounds		
6	499	61.6	508	6	BD265239	BD265239	Compounds		
7	499	61.6	508	6	AR401213	AR401213	Sequence		
8	499	61.6	508	6	AR401225	AR401225	Sequence		
9	499	61.6	508	6	AX192666	AX192666	Sequence		
10	499	61.6	508	6	AX192678	AX192678	Sequence		
11	471.6	58.2	630	10	AF353993	AF353993	Mus muscu		
12	436.8	53.9	440	6	BD076181	BD076181	5' EST of		
c 13	430.6	53.2	193660	2	AC102262	AC102262	Mus muscu		
14	425.2	52.5	455	6	BD076249	BD076249	5' EST of		
15	411.4	50.8	487	6	AX892343	AX892343	Sequence		
16	411.4	50.8	487	6	BD027876	BD027876	Sequence		
c 17	354	43.7	183425	9	AC097064	AC097064	Homo sapi		
c 18	354	43.7	239704	9	AC099791	AC099791	Homo sapi		
c 19	325.4	40.2	228458	2	AC097670	AC097670	Rattus no		
20	299	36.9	176056	10	AC073437	AC073437	Mus muscu		
21	299	36.9	196421	10	AL672100	AL672100	Mouse DNA		
22	276	34.1	185576	2	AC025691	AC025691	Homo sapi		
c 23	250.4	30.9	231150	2	AC114195	AC114195	Rattus no		
24	244.4	30.2	157999	2	AC117088	AC117088	Rattus no		
25	209.8	25.9	277191	2	AC109077	AC109077	Rattus no		
26	186.2	23.0	167627	9	AC079382	AC079382	Homo sapi		
27	145.6	18.0	178068	2	AC142046	AC142046	Rattus no		
c 28	114	14.1	129705	2	AC133258	AC133258	Rattus no		
29	114	14.1	239113	2	AC094034	AC094034	Rattus no		
30	114	14.1	324462	2	AC137263	AC137263	Rattus no		
c 31	107.6	13.3	145871	2	AC143611	AC143611	Macaca mu		
32	99.8	12.3	298	6	E25986	E25986	Blastocyst		
33	99.2	12.2	240950	2	AC098287	AC098287	Rattus no		

c	34	97.8	12.1	131215	10	AL671140	AL671140 Mouse DNA
c	35	73.2	9.0	198395	2	AC101700	AC101700 Mus muscu
	36	64.6	8.0	128444	2	AC019924	AC019924 Drosophil
	37	64.6	8.0	132637	2	AC006092	AC006092 Drosophil
c	38	64.6	8.0	149592	2	AC005718	AC005718 Drosophil
c	39	64.6	8.0	179139	3	AC099307	AC099307 Drosophil
c	40	64.6	8.0	188633	3	AC007175	AC007175 Drosophil
c	41	64.6	8.0	305150	3	AE003453	AE003453 Drosophil
	42	64.4	8.0	1052	3	AY061343	AY061343 Drosophil
	43	58.2	7.2	155623	5	AL929239	AL929239 Zebrafish
	44	58.2	7.2	239459	2	BX322568	BX322568 Danio rer
c	45	54.6	6.7	214455	2	AC118451	AC118451 Rattus no

# ALIGNMENTS

## RESULT 1

BD243134

LOCUS BD243134 810 bp DNA linear PAT 17-JUL-2003

DEFINITION 6-Protein-bound receptor-like protein, polynucleotide encoded by it, and method of use thereof.

ACCESSION BD243134

VERSION BD243134.1 GI:33052904

KEYWORDS JP 2002527064-A/1.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 810)

AUTHORS Ozenberger, B.A., Kajkowski, E.M. and Lo, C.H.F.

TITLE 6-Protein-bound receptor-like protein, polynucleotide encoded by it, and method of use thereof

JOURNAL Patent: JP 2002527064-A 1 27-AUG-2002;

AMERICAN HOME PRODUCTS CORP

COMMENT OS Homo sapiens (human)

PN JP 2002527064-A/1

PD 27-AUG-2002

PF 13-OCT-1999 JP 2000576015

PR 13-OCT-1998 US 60/104104

PI BRADLEY ALTON OZENBERGER, EILEEN MARIE KAJKOWSKI, CHING HSIUNG

PI FREDERICK LO

PC C12N15/09, A61K45/00, A61P43/00, C07K14/705, C12N1/15, C12N1/19, PC C12N1/21,

PC C12N5/10, C12Q1/02, C12Q1/68, G01N33/15, G01N33/50, G01N33/53, PC G01N33/566,

PC C12N15/00, C12N5/00

CC 6-Protein-bound receptor-like protein, polynucleotide encoded by it, and

CC method of use thereof

FH Key Location/Qualifiers

FT source 1. .810

FT /organism='Homo sapiens (human)'.

FEATURES Location/Qualifiers

source 1. .810

/organism="Homo sapiens"

/mol\_type="genomic DNA"



## ORIGIN

Qy	1	ATGCATATTTTAAAGGGTCTCCCAATGTGATTCCACGGGCTCACGGGCAGAAGAACACG	60
Db	1	ATGCATATTTTAAAGGGTCTCCCAATGTGATTCCACGGGCTCACGGGCAGAAGAACACG	60
Qy	61	CGAAGAGACGGAAGTGGCCTCTATCCTATGCGAGGTCCCTTTAAGAACCTCGCCCTGTTG	120
Db	61	CGAAGAGACGGAAGTGGCCTCTATCCTATGCGAGGTCCCTTTAAGAACCTCGCCCTGTTG	120
Qy	121	CCCTTCTCCCTCCCGCTCCTGGGCGGAGGCGGAAGCGGAAGTGGCGAGAAAGTGTGGTC	180
Db	121	CCCTTCTCCCTCCCGCTCCTGGGCGGAGGCGGAAGCGGAAGTGGCGAGAAAGTGTGGTC	180
Qy	181	TCCAAGATGGCGGCCGCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGGCCAGA	240
Db	181	TCCAAGATGGCGGCCGCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGGCCAGA	240
Qy	241	CTCGTTGGTGTCTGTGGTTTCGTCTCAGTCACTACAGGACCTGGGGGGCTGTTGCCACC	300
Db	241	CTCGTTGGTGTCTGTGGTTTCGTCTCAGTCACTACAGGACCTGGGGGGCTGTTGCCACC	300
Qy	301	TCCGCCGGGGGCGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATATTTGT	360
Db	301	TCCGCCGGGGGCGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATATTTGT	360
Qy	361	AAAGATCCAAAAATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCT	420
Db	361	AAAGATCCAAAAATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCT	420
Qy	421	CATGTTTCCTGTTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACA	480
Db	421	CATGTTTCCTGTTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACA	480
Qy	481	CATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGC	540
Db	481	CATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGC	540
Qy	541	TATTCCTACAAAGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGA	600
Db	541	TATTCCTACAAAGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGA	600
Qy	601	TTTACCTTGGATACCCCTGCTTTGGGTTTGTTAAAGTTTTGCACTGTAGGGTTTTGTGGA	660
Db	601	TTTACCTTGGATACCCCTGCTTTGGGTTTGTTAAAGTTTTGCACTGTAGGGTTTTGTGGA	660
Qy	661	ATTGGGAGCCTAATTGATTTCAATCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGA	720
Db	661	ATTGGGAGCCTAATTGATTTCAATCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGA	720
Qy	721	AGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAA	780

Db 721 AGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAA 780

QY 781 ACATTTAGAAAAACGCAATTATATCCATAA 810  
 |||

Db 781 ACATTTAGAAAAACGCAATTATATCCATAA 810

RESULT 2

AF353990

LOCUS AF353990 1246 bp mRNA linear PRI 29-MAY-2001

DEFINITION Homo sapiens beta-amyloid binding protein precursor (BBP) mRNA,  
 complete cds.

ACCESSION AF353990

VERSION AF353990.1 GI:13625458

KEYWORDS .

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1246)

AUTHORS Kajkowski,E.M., Lo,C.F., Ning,X., Walker,S., Sofia,H.J., Wang,W.,  
 Edris,W., Chanda,P., Wagner,E., Vile,S., Ryan,K.,  
 McHendry-Rinde,B., Smith,S.C., Wood,A., Rhodes,K.J., Kennedy,J.D.,  
 Bard,J., Jacobsen,J.S. and Ozenberger,B.A.

TITLE beta -Amyloid peptide-induced apoptosis regulated by a novel  
 protein containing a g protein activation module

JOURNAL J. Biol. Chem. 276 (22), 18748-18756 (2001)

MEDLINE 21276355

PUBMED 11278849

REFERENCE 2 (bases 1 to 1246)

AUTHORS Ozenberger,B.A., Kajkowski,E., Jacobsen,J.S., Bard,J. and Walker,S.

TITLE Direct Submission

JOURNAL Submitted (27-FEB-2001) Wyeth Neuroscience, CN 8000, Princeton, NJ  
 08543, USA

FEATURES Location/Qualifiers

source

1. .1246  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /chromosome="1"

gene

1. .1246  
 /gene="BBP"

CDS

304. .927  
 /gene="BBP"  
 /note="membrane-associated glycoprotein"  
 /codon\_start=1  
 /product="beta-amyloid binding protein precursor"  
 /protein\_id="AAK35064.1"  
 /db\_xref="GI:13625459"  
 /translation="MAAAWPSGPSAPEAVTARLVGVLFVSVTTGPWGAVATSAGGEE  
 SLKCEDLVKGQYICKDPKINDATQEPVNCTNYTAHVSCFPAPNITCKDSSGNETHFTG  
 NEVGFFKPISCRNVNGYSYKVAVALSLFLGWLGADRFYLGYPALGLLKFCYVGFCEGIG  
 SLIDFILISMQIVGPSDGSSYIIDYYGTRLRLSITNETFRKTQLYP"

sig\_peptide

304. .414  
 /gene="BBP"

misc\_feature

646. .711  
 /gene="BBP"

misc\_feature            /note="Region: transmembrane domain"  
751. .825  
/gene="BBP"  
/note="Region: transmembrane domain"

ORIGIN

Query Match            100.0%;    Score 810;    DB 9;    Length 1246;  
Best Local Similarity   100.0%;    Pred. No. 2.4e-206;  
Matches   810;    Conservative    0;    Mismatches    0;    Indels    0;    Gaps    0;

```
Qy      1 ATGCATATTTTAAAGGGTCTCCCAATGTGATTCCACGGGCTCACGGGCAGAAGAACACG 60
        |||
Db     118 ATGCATATTTTAAAGGGTCTCCCAATGTGATTCCACGGGCTCACGGGCAGAAGAACACG 177

Qy      61 CGAAGAGACGGAACCTGGCCTCTATCCTATGCGAGGTCCCTTTAAGAACCTCGCCCTGTTG 120
        |||
Db     178 CGAAGAGACGGAACCTGGCCTCTATCCTATGCGAGGTCCCTTTAAGAACCTCGCCCTGTTG 237

Qy     121 CCCTTCTCCCTCCCGCTCCTGGGCGGAGGCGGAAGCGGAAGTGGCGAGAAAGTGTGGTC 180
        |||
Db     238 CCCTTCTCCCTCCCGCTCCTGGGCGGAGGCGGAAGCGGAAGTGGCGAGAAAGTGTGGTC 297

Qy     181 TCCAAGATGGCGGCCGCCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGGCCAGA 240
        |||
Db     298 TCCAAGATGGCGGCCGCCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGGCCAGA 357

Qy     241 CTCGTTGGTGTCTGTGGTTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGCCACC 300
        |||
Db     358 CTCGTTGGTGTCTGTGGTTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGCCACC 417

Qy     301 TCCGCCGGGGGCGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGT 360
        |||
Db     418 TCCGCCGGGGGCGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGT 477

Qy     361 AAAGATCCAAAAATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACACAGCT 420
        |||
Db     478 AAAGATCCAAAAATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACACAGCT 537

Qy     421 CATGTTTCCTGTTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACA 480
        |||
Db     538 CATGTTTCCTGTTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACA 597

Qy     481 CATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGC 540
        |||
Db     598 CATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGC 657

Qy     541 TATTCCCTACAAAGTGGCAGTCGCATTGTCTCTTTTTCTTGATGGTTGGGAGCAGATCGA 600
        |||
Db     658 TATTCCCTACAAAGTGGCAGTCGCATTGTCTCTTTTTCTTGATGGTTGGGAGCAGATCGA 717

Qy     601 TTTTACCTTGATACCTGCTTTGGGTTTGTTAAAGTTTTGCACTGTAGGGTTTTGTGGA 660
        |||
Db     718 TTTTACCTTGATACCTGCTTTGGGTTTGTTAAAGTTTTGCACTGTAGGGTTTTGTGGA 777

Qy     661 ATTGGGAGCCTAATTGATTTCATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGA 720
        |||
Db     778 ATTGGGAGCCTAATTGATTTCATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGA 837
```

Qy 721 AGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAA 780  
 |||  
 Db 838 AGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAA 897  
 |||  
 Qy 781 ACATTTAGAAAAACGCAATTATATCCATAA 810  
 |||  
 Db 898 ACATTTAGAAAAACGCAATTATATCCATAA 927

RESULT 3

BC029486

LOCUS BC029486 984 bp mRNA linear PRI 06-OCT-2003

DEFINITION Homo sapiens beta-amyloid binding protein precursor, mRNA (cDNA clone MGC:32941 IMAGE:5271098), complete cds.

ACCESSION BC029486

VERSION BC029486.1 GI:20809565

KEYWORDS MGC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 984)

AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

MEDLINE 22388257

PUBMED 12477932

REFERENCE 2 (bases 1 to 984)

AUTHORS Strausberg,R.

TITLE Direct Submission

JOURNAL Submitted (01-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www-shgc.stanford.edu>  
 Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAK Plate: 48 Row: b Column: 24  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 17738309.

```

FEATURES             Location/Qualifiers
     source            1. .984
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="MGC:32941 IMAGE:5271098"
                        /tissue_type="Testis"
                        /clone_lib="NIH_MGC_97"
                        /lab_host="DH10B"
                        /note="Vector: pBluescript"
     gene              1. .984
                        /gene="BBP"
                        /db_xref="LocusID:83941"
     CDS                22. .645
                        /codon_start=1
                        /product="beta-amyloid binding protein precursor"
                        /protein_id="AAH29486.1"
                        /db_xref="GI:20809566"
                        /db_xref="LocusID:83941"
                        /translation="MAAAWPSPGSAPEAVTARLVGVLFVSVTTGPGWAVATSAGGEE
                        SLKCEDLKVGQYICKDPKINDATQEPVNCTNYTAHVSCFPAPNITCKDSSGNETHFTG
                        NEVGFFKPISCRNVNGYSYKVAVALSLFLGWLGADREFYLGYPALGLLKFCFTVGFCGIG
                        SLIDFILISMQIVGPSDGSSYIIDYYGTRLRLSITNETFRKTQLYP"
     misc_feature      385. .525
                        /note="XynA; Region: Predicted membrane protein [Function
                        unknown]"
                        /db_xref="CDD:COG2314"

```

#### ORIGIN

Query Match 79.6%; Score 645; DB 9; Length 984;

Best Local Similarity 100.0%; Pred. No. 4.8e-162;

Matches 645; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      166 GAGAAAGTGTCTGGTCTCCAAGATGGCGGCCGCCTGGCCGTCTGGTCCGTCTGCTCCGGAG 225
          |||
Db      1   GAGAAAGTGTCTGGTCTCCAAGATGGCGGCCGCCTGGCCGTCTGGTCCGTCTGCTCCGGAG 60

Qy      226 GCCGTGACGGCCAGACTCGTTGGTGTCTGTGGTTTCGTCTCAGTCACTACAGGACCCTGG 285
          |||
Db      61 GCCGTGACGGCCAGACTCGTTGGTGTCTGTGGTTTCGTCTCAGTCACTACAGGACCCTGG 120

Qy      286 GGGGCTGTTGCCACCTCCGCCGGGGGCGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTG 345

```

```

      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 GGGGCTGTTGCCACCTCCGCCGGGGGCGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTG 180
Qy      346 GGACAATATATTTGTAAAGATCCAAAAATAAATGACGCTACGCAAGAACCAGTTAACTGT 405
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 GGACAATATATTTGTAAAGATCCAAAAATAAATGACGCTACGCAAGAACCAGTTAACTGT 240
Qy      406 ACAAACTACACAGCTCATGTTTCCTGTTTTCAGCACCCCAACATAACTTGTAAGGATTCC 465
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 ACAAACTACACAGCTCATGTTTCCTGTTTTCAGCACCCCAACATAACTTGTAAGGATTCC 300
Qy      466 AGTGGCAATGAAACACATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGC 525
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 AGTGGCAATGAAACACATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGC 360
Qy      526 CGAAATGTAAATGGCTATTTCCTACAAAGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGG 585
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 CGAAATGTAAATGGCTATTTCCTACAAAGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGG 420
Qy      586 TTGGGAGCAGATCGATTTTACCTTGGATACCCTGCTTTGGGTTTGTTAAAGTTTGCAGT 645
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 TTGGGAGCAGATCGATTTTACCTTGGATACCCTGCTTTGGGTTTGTTAAAGTTTGCAGT 480
Qy      646 GTAGGGTTTTGTGGAATTGGGAGCCTAATTGATTTTCAATTCATTCAATGCAGATTGTT 705
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 GTAGGGTTTTGTGGAATTGGGAGCCTAATTGATTTTCAATTCATTCAATGCAGATTGTT 540
Qy      706 GGACCTTCAGATGGAAGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTG 765
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 GGACCTTCAGATGGAAGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTG 600
Qy      766 AGTATTACTAATGAAACATTTAGAAAAACGCAATTATATCCATAA 810
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 AGTATTACTAATGAAACATTTAGAAAAACGCAATTATATCCATAA 645

```

#### RESULT 4

BD139411

LOCUS BD139411 970 bp DNA linear PAT 18-SEP-2002

DEFINITION Extended cDNA of secretory protein.

ACCESSION BD139411

VERSION BD139411.1 GI:23234356

KEYWORDS JP 2002508182-A/163.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 970)

AUTHORS Bougueleret, L., Duclert, A. and Edwards, J.B.D.M.

TITLE Extended cDNA of secretory protein

JOURNAL Patent: JP 2002508182-A 163 19-MAR-2002;

GENSET

COMMENT OS Homo sapiens (human)

PN JP 2002508182-A/163

PD 19-MAR-2002

PF 17-DEC-1998 JP 2000539136

PR 17-DEC-1997 US 60/069957, 09-FEB-1998 US 60/074121 PR

13-APR-1998 US 60/081563,10-AUG-1998 US 60/096116 PI LYDIE  
BOUGUELERET,AYMERIC DUCLERT,JEAN BAPTISTE DUMAS MILNE PI EDWARDS  
PC C12N15/09,C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19, PC  
C12N1/21,

PC C12N5/10,C12P21/02,C12Q1/68,C12N15/00,C12N5/00,C12N15/00 CC

Von Heijne matrix

CC score 5.5

CC seq LVGVLFVSVTTG/PW

FH Key Location/Qualifiers

FT CDS 12. .497

FT sig\_peptide 12. .104

FT polyA\_signal 935. .940

FT polyA\_site 955. .967.

FEATURES Location/Qualifiers  
source 1. .970  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

#### ORIGIN

Query Match 74.4%; Score 602.8; DB 6; Length 970;  
Best Local Similarity 98.4%; Pred. No. 1e-150;  
Matches 624; Conservative 5; Mismatches 3; Indels 2; Gaps 2;

Qy	177	GGTCTCCAAGATGGCGGCCGCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGGC	236
Db	2	GGTCTCCAAGATGGCGGCCGCTGGCCGTCTGGTCCGKCTGCTCCGGAGGCCGTGACGGC	61
Qy	237	CAGACTCGTTGGTGTCTGTGGTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGC	296
Db	62	CAGACTCGTTGGTGTCTGTGGTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGC	121
Qy	297	CACCTCCGCCGGGGGCGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATAT	356
Db	122	CACCTCCGCCGGGGGCGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATAT	181
Qy	357	TTGTAAAGATCCAAAAATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACAC	416
Db	182	TTGTAAAGATCCAAAAATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACAC	241
Qy	417	AGCTCATGTTTCCTGTTTTCCAGCACCCAACATAACTTGTAAAGGATTCCAGTGGCAATGA	476
Db	242	AGCTCATGTTTCCTGTTTTCCAGCACCCAACATAACTTGTAAAGGATTCCAGTGGCAATGA	301
Qy	477	AACACATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAA	536
Db	302	AACACATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAA	361
Qy	537	TGGCTATTCCTACAAAGTGGCAGTCGCATTGTCTCTTTTCTTGGATGGTTGGGAGCAGA	596
Db	362	TGGCTATTCCTACAATG-AGCAGTCGCA-TGTCTCTTTTCTTGGATGGTTGGGAGCAGA	419
Qy	597	TCGATTTTACCTTGGATACCCTGCTTTGGGTTTGTAAAGTTTGGCACTGTAGGGTTTGG	656
Db	420	TCGATTTTACCTTGGATACCCTGCTTTGGGTTTGTAAABTTYGCACTGTAGGGTTTKG	479
Qy	657	TGGAATTGGGAGCCTAATTGATTTTATTCTTATTTCAATGCAGATTGTTGGACCTTCAGA	716

```

      |||||||:|||||
Db      480 TGG AATTGGGAGCCTAATTGATTTCA TYCTATTTCAATGCAGATTGTTGGACCTTCAA 539

Qy      717 TGG AAGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAA 776
      |||||||
Db      540 TGG AAGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAA 599

Qy      777 TGAAACATTTAGAAAAACGCAATTATATCCATAA 810
      |||||||
Db      600 TGAAACATTTAGAAAAACGCAATTATATCCATAA 633

```

RESULT 5

BD265227

LOCUS BD265227 508 bp DNA linear PAT 17-JUL-2003

DEFINITION Compounds for immunotherapy and diagnosis of colonic cancer and method of using the same.

ACCESSION BD265227

VERSION BD265227.1 GI:33074995

KEYWORDS JP 2002533082-A/225.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 508)

AUTHORS Xu, J., Lodes, M.J., Secrist, H., Benson, D.R., Meagher, M.J., Stolk, J., Wang, T. and Yuqiu, J.

TITLE Compounds for immunotherapy and diagnosis of colonic cancer and method of using the same

JOURNAL Patent: JP 2002533082-A 225 08-OCT-2002; CORIXA CORP

COMMENT OS Homo sapiens (human)

PN JP 2002533082-A/225

PD 08-OCT-2002

PF 23-DEC-1999 JP 2000589697

PR 23-DEC-1998 US 09/221298, 02-JUL-1999 US 09/347496 PR

22-SEP-1999 US 09/401064, 19-NOV-1999 US 09/444242 PR

02-DEC-1999 US 09/454150

PI JIANGCHUN XU, MICHAEL J LODES, HEATHER SECRIST, DARIN R BENSON,

PI MADELEINE JOY MEAGHER, JOHN STOLK, TONGTONG WANG, JIANG YUQIU PC C12N15/09, A61K31/711, A61K35/14, A61K38/00, A61K39/00, A61K39/395, PC A61K39/395,

PC A61P35/00, C07K14/47, C07K16/18, C07K19/00, C12N1/15, C12N1/19, PC C12N1/21,

PC

C12N5/06, C12N5/10, C12Q1/02, C12Q1/68, G01N33/53, G01N33/53, G01N33/ PC 566,

PC G01N33/574, G01N33/577, G01N33/58, C12N15/00, C12N5/00, C12N5/00,

PC A61K37/02

CC Compounds for immunotherapy and diagnosis of colonic cancer

CC and method of

CC using the same

FH Key Location/Qualifiers

FT source 1. .508

FT /organism='Homo sapiens (human)'. .

FEATURES Location/Qualifiers

source 1. .508



/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 61.6%; Score 499; DB 6; Length 508;  
Best Local Similarity 100.0%; Pred. No. 6.9e-123;  
Matches 499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      312 CGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 371
          |||
Db       1  CGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 60

Qy      372 AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCTCATGTTTCCTG 431
          |||
Db      61  AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCTCATGTTTCCTG 120

Qy      432 TTTTCCAGCACCCCAACATAAATTGTAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 491
          |||
Db     121  TTTTCCAGCACCCCAACATAAATTGTAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 180

Qy      492 GAACGAAGTTGGTTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCTCTACAA 551
          |||
Db     181  GAACGAAGTTGGTTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCTCTACAA 240

Qy      552 AGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG 611
          |||
Db     241  AGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG 300

Qy      612 ATACCCTGCTTTGGGTTTGTTAAAGTTTTGCACTGTAGGGTTTTGTGGAATTGGGAGCCT 671
          |||
Db     301  ATACCCTGCTTTGGGTTTGTTAAAGTTTTGCACTGTAGGGTTTTGTGGAATTGGGAGCCT 360

Qy      672 AATTGATTTCATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 731
          |||
Db     361  AATTGATTTCATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 420

Qy      732 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 791
          |||
Db     421  TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 480

Qy      792 AACGCAATTATATCCATAA 810
          |||
Db     481  AACGCAATTATATCCATAA 499
```

RESULT 6

BD265239

LOCUS BD265239 508 bp DNA linear PAT 17-JUL-2003

DEFINITION Compounds for immunotherapy and diagnosis of colonic cancer and  
method of using the same.

ACCESSION BD265239

VERSION BD265239.1 GI:33075007

KEYWORDS JP 2002533082-A/237.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 508)

AUTHORS Xu, J., Lodes, M. J., Secrist, H., Benson, D. R., Meagher, M. J., Stolk, J., Wang, T. and Yuqiu, J.

TITLE Compounds for immunotherapy and diagnosis of colonic cancer and method of using the same

JOURNAL Patent: JP 2002533082-A 237 08-OCT-2002;  
CORIXA CORP

COMMENT OS Homo sapiens (human)  
PN JP 2002533082-A/237  
PD 08-OCT-2002  
PF 23-DEC-1999 JP 2000589697  
PR 23-DEC-1998 US 09/221298, 02-JUL-1999 US 09/347496 PR  
22-SEP-1999 US 09/401064, 19-NOV-1999 US 09/444242 PR  
02-DEC-1999 US 09/454150  
PI JIANGCHUN XU, MICHAEL J LODES, HEATHER SECRIST, DARIN R BENSON,  
PI MADELEINE JOY MEAGHER, JOHN STOLK, TONGTONG WANG, JIANG YUQIU PC  
C12N15/09, A61K31/711, A61K35/14, A61K38/00, A61K39/00, A61K39/395, PC  
A61K39/395,  
PC A61P35/00, C07K14/47, C07K16/18, C07K19/00, C12N1/15, C12N1/19, PC  
C12N1/21,  
PC  
C12N5/06, C12N5/10, C12Q1/02, C12Q1/68, G01N33/53, G01N33/53, G01N33/ PC  
566,  
PC G01N33/574, G01N33/577, G01N33/58, C12N15/00, C12N5/00, C12N5/00,  
PC A61K37/02  
CC Compounds for immunotherapy and diagnosis of colonic cancer  
CC and method of  
CC using the same  
FH Key Location/Qualifiers  
FT source 1. .508  
FT /organism='Homo sapiens (human)'.  
FT

FEATURES Location/Qualifiers  
source 1. .508  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

# ORIGIN

Query Match 61.6%; Score 499; DB 6; Length 508;  
Best Local Similarity 100.0%; Pred. No. 6.9e-123;  
Matches 499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	312	CGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA	371
Db	1	CGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA	60

---

Qy	372	AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCTCATGTTTCCTG	431
Db	61	AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCTCATGTTTCCTG	120

---

Qy	432	TTTTCCAGCACCCAACATAACTTGTAAGGATTCAGTGGCAATGAAACACATTTACTGG	491
Db	121	TTTTCCAGCACCCAACATAACTTGTAAGGATTCAGTGGCAATGAAACACATTTACTGG	180

---

Qy	492	GAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCCTACAA	551

```

Db      181 GAACGAAGTTGGTTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCTTACAA 240
Qy      552 AGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG 611
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 AGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG 300
Qy      612 ATACCCTGCTTTGGGTTTGTAAAGTTTGCAGTGTAGGGTTTGTGGAATTGGGAGCCT 671
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 ATACCCTGCTTTGGGTTTGTAAAGTTTGCAGTGTAGGGTTTGTGGAATTGGGAGCCT 360
Qy      672 AATTGATTTTATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 731
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 AATTGATTTTATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 420
Qy      732 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 791
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 480
Qy      792 AACGCAATTATATCCATAA 810
        ||||||||||||||||
Db      481 AACGCAATTATATCCATAA 499

```

RESULT 7

AR401213

LOCUS AR401213 508 bp DNA linear PAT 18-DEC-2003

DEFINITION Sequence 233 from patent US 6623923.

ACCESSION AR401213

VERSION AR401213.1 GI:40148513

KEYWORDS .

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 508)

AUTHORS Xu,J., Lodes,M.J., Secrist,H., Meagher,M.J., Stolk,J., Benson,D.R.  
and Wang,T.

TITLE Compounds for immunotherapy and diagnosis of colon cancer and  
methods for their use

JOURNAL Patent: US 6623923-A 233 23-SEP-2003;

FEATURES Location/Qualifiers

source 1. .508

/organism="unknown"

/mol\_type="genomic DNA"

ORIGIN

Query Match 61.6%; Score 499; DB 6; Length 508;

Best Local Similarity 100.0%; Pred. No. 6.9e-123;

Matches 499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      312 CGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 371
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 60
Qy      372 AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACACAGCTCATGTTTCCTG 431
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACACAGCTCATGTTTCCTG 120

```

Qy 432 TTTTCCAGCACCCAACATAAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 491  
 |||  
 Db 121 TTTTCCAGCACCCAACATAAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 180  
 Qy 492 GAACGAAGTTGGTTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCTTACAA 551  
 |||  
 Db 181 GAACGAAGTTGGTTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCTTACAA 240  
 Qy 552 AGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG 611  
 |||  
 Db 241 AGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG 300  
 Qy 612 ATACCCTGCTTTGGGTTTGTAAAGTTTTGCACTGTAGGGTTTGTGGAATTGGGAGCCT 671  
 |||  
 Db 301 ATACCCTGCTTTGGGTTTGTAAAGTTTTGCACTGTAGGGTTTGTGGAATTGGGAGCCT 360  
 Qy 672 AATTGATTTTCATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 731  
 |||  
 Db 361 AATTGATTTTCATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 420  
 Qy 732 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 791  
 |||  
 Db 421 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 480  
 Qy 792 AACGCAATTATATCCATAA 810  
 |||  
 Db 481 AACGCAATTATATCCATAA 499

# RESULT 8

AR401225

LOCUS AR401225 508 bp DNA linear PAT 18-DEC-2003

DEFINITION Sequence 245 from patent US 6623923.

ACCESSION AR401225

VERSION AR401225.1 GI:40148525

KEYWORDS .

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 508)

AUTHORS Xu,J., Lodes,M.J., Secrist,H., Meagher,M.J., Stolk,J., Benson,D.R.  
 and Wang,T.

TITLE Compounds for immunotherapy and diagnosis of colon cancer and  
 methods for their use

JOURNAL Patent: US 6623923-A 245 23-SEP-2003;

FEATURES Location/Qualifiers

source 1..508

/organism="unknown"

/mol\_type="genomic DNA"

## ORIGIN

Query Match 61.6%; Score 499; DB 6; Length 508;

Best Local Similarity 100.0%; Pred. No. 6.9e-123;

Matches 499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 312 CGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 371  
 |||

Db 1 CGAGGAGTCGCTTAAGT GCGAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 60

Qy 372 AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCTCATGTTTCCTG 431  
 |||

Db 61 AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCTCATGTTTCCTG 120

Qy 432 TTTTCCAGCACCCAACATAAATTGTAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 491  
 |||

Db 121 TTTTCCAGCACCCAACATAAATTGTAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 180

Qy 492 GAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCTCTACAA 551  
 |||

Db 181 GAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCTCTACAA 240

Qy 552 AGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG 611  
 |||

Db 241 AGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG 300

Qy 612 ATACCCTGCTTTGGGTTTGTAAAGTTTTGCACTGTAGGGTTTTGTGGAATTGGGAGCCT 671  
 |||

Db 301 ATACCCTGCTTTGGGTTTGTAAAGTTTTGCACTGTAGGGTTTTGTGGAATTGGGAGCCT 360

Qy 672 AATTGATTTCATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 731  
 |||

Db 361 AATTGATTTCATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 420

Qy 732 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 791  
 |||

Db 421 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 480

Qy 792 AACGCAATTATATCCATAA 810  
 |||

Db 481 AACGCAATTATATCCATAA 499

# RESULT 9

AX192666

LOCUS AX192666 508 bp DNA linear PAT 15-AUG-2001

DEFINITION Sequence 233 from Patent WO0149716.

ACCESSION AX192666

VERSION AX192666.1 GI:15210622

KEYWORDS .

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Xu,J., Lodes,M.J., Secrist,H., Benson,D.R., Meagher,M.J.,  
 Stolk,J.A., King,G.E., Wang,T. and Jiang,Y.

TITLE Compounds for immunotherapy and diagnosis of colon cancer and  
 methods for their use

JOURNAL Patent: WO 0149716-A 233 12-JUL-2001;  
 CORIXA CORPORATION (US)

FEATURES Location/Qualifiers

source 1. .508

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

ORIGIN

Query Match 61.6%; Score 499; DB 6; Length 508;  
Best Local Similarity 100.0%; Pred. No. 6.9e-123;  
Matches 499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      312 CGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 371
          |||
Db       1  CGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 60

Qy      372 AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCTCATGTTTCCTG 431
          |||
Db       61 AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCTCATGTTTCCTG 120

Qy      432 TTTTCCAGCACCCAACATAAATTGTAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 491
          |||
Db      121 TTTTCCAGCACCCAACATAAATTGTAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 180

Qy      492 GAACGAAGTTGGTTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCTCTACAA 551
          |||
Db      181 GAACGAAGTTGGTTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCTCTACAA 240

Qy      552 AGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG 611
          |||
Db      241 AGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG 300

Qy      612 ATACCCTGCTTTGGGTTTGTAAAGTTTTGCACTGTAGGGTTTTGTGGAATTGGGAGCCT 671
          |||
Db      301 ATACCCTGCTTTGGGTTTGTAAAGTTTTGCACTGTAGGGTTTTGTGGAATTGGGAGCCT 360

Qy      672 AATTGATTTTCATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 731
          |||
Db      361 AATTGATTTTCATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 420

Qy      732 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 791
          |||
Db      421 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 480

Qy      792 AACGCAATTATATCCATAA 810
          |||
Db      481 AACGCAATTATATCCATAA 499
```

RESULT 10

AX192678

LOCUS AX192678 508 bp DNA linear PAT 15-AUG-2001

DEFINITION Sequence 245 from Patent WO0149716.

ACCESSION AX192678

VERSION AX192678.1 GI:15210634

KEYWORDS .

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Xu,J., Lodes,M.J., Secrist,H., Benson,D.R., Meagher,M.J.,

TITLE           Stolk,J.A., King,G.E., Wang,T. and Jiang,Y.  
                   Compounds for immunotherapy and diagnosis of colon cancer and  
                   methods for their use  
 JOURNAL        Patent: WO 0149716-A 245 12-JUL-2001;  
                   CORIXA CORPORATION (US)  
 FEATURES        Location/Qualifiers  
                   source           1. .508  
                                   /organism="Homo sapiens"  
                                   /mol\_type="unassigned DNA"  
                                   /db\_xref="taxon:9606"

ORIGIN

Query Match                   61.6%;   Score 499;   DB 6;   Length 508;  
 Best Local Similarity       100.0%;   Pred. No. 6.9e-123;  
 Matches 499;   Conservative   0;   Mismatches   0;   Indels    0;   Gaps    0;

```

Qy      312 CGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 371
          |||
Db      1   CGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 60

Qy      372 AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCTCATGTTTCCTG 431
          |||
Db      61  AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCTCATGTTTCCTG 120

Qy      432 TTTTCCAGCACCCAACATAAATTGTAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 491
          |||
Db      121 TTTTCCAGCACCCAACATAAATTGTAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 180

Qy      492 GAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCCTACAA 551
          |||
Db      181 GAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCCTACAA 240

Qy      552 AGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG 611
          |||
Db      241 AGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG 300

Qy      612 ATACCCTGCTTTGGGTTTGTTAAAGTTTTGCACTGTAGGGTTTTGTGGAATTGGGAGCCT 671
          |||
Db      301 ATACCCTGCTTTGGGTTTGTTAAAGTTTTGCACTGTAGGGTTTTGTGGAATTGGGAGCCT 360

Qy      672 AATTGATTTCATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 731
          |||
Db      361 AATTGATTTCATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 420

Qy      732 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 791
          |||
Db      421 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 480

Qy      792 AACGCAATTATATCCATAA 810
          |||
Db      481 AACGCAATTATATCCATAA 499
  
```

RESULT 11

AF353993

LOCUS           AF353993                   630 bp   mRNA   linear   ROD 29-MAY-2001  
 DEFINITION   Mus musculus beta-amyloid binding protein (Bbp) mRNA, complete cds.







COMMENT OS Homo sapiens (human)  
 PN JP 2001512011-A/129  
 PD 21-AUG-2001  
 PF 31-JUL-1998 JP 2000505289  
 PR 01-AUG-1997 US 08/905135  
 PI JEAN BAPTISTE DUMAS MILNE EDWARDS, AYMERIC DUCLERT, BRUNO PI  
 LACROIX  
 PC C12N15/09, C12N15/09, C07K14/47, C12Q1/68, C12N15/00, C12N15/00 CC  
 blastn  
 CC identity 97  
 CC region 113. .315  
 CC id AA143062  
 CC est  
 CC blastn  
 CC identity 99  
 CC region 304. .411  
 CC id AA143062  
 CC est  
 CC blastn  
 CC identity 97  
 CC region 43. .120  
 CC id AA143062  
 CC est  
 CC blastn  
 CC identity 97  
 CC region 44. .317  
 CC id HUM172D06B  
 CC est  
 CC blastn  
 CC identity 100  
 CC region 340. .410  
 CC id HUM172D06B  
 CC est  
 CC blastn  
 CC identity 94  
 CC region 8. .46  
 CC id HUM172D06B  
 CC est  
 CC blastn  
 CC identity 100  
 CC region 125. .414  
 CC id N47594  
 CC est  
 CC blastn  
 CC identity 100  
 CC region 49. .119  
 CC id N47594  
 CC est  
 CC blastn  
 CC identity 98  
 CC region 45. .385  
 CC id HUM159G08B  
 CC est  
 CC blastn  
 CC identity 95  
 CC region 1. .47  
 CC id HUM159G08B

```

CC     est
CC     blastn
CC     identity 98
CC     region 92. .316
CC     id N34957
CC     est
CC     blastn
CC     identity 100
CC     region 30. .97
CC     id N34957
CC     est
CC     blastn
CC     identity 91
CC     region 312. .379
CC     id N34957
CC     est
CC     Von Heijne matrix
CC     score 8.7
CC     seq AVALSFLGWLGA/DR
FH     Key                      Location/Qualifiers
FT     misc_feature             143. .345
FT     misc_feature             335. .442
FT     misc_feature             72. .149
FT     misc_feature             72. .345
FT     misc_feature             372. .442
FT     misc_feature             35. .73
FT     misc_feature             153. .442
FT     misc_feature             77. .147
FT     misc_feature             72. .412
FT     misc_feature             27. .73
FT     misc_feature             143. .367
FT     misc_feature             80. .147
FT     misc_feature             362. .429
FT     sig_peptide              24. .431.

```

```

FEATURES             Location/Qualifiers
     source           1. .440
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"

```

# ORIGIN

```

Query Match          53.9%;  Score 436.8;  DB 6;  Length 440;
Best Local Similarity 99.5%;  Pred. No. 3.4e-106;
Matches 436;  Conservative 2;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

Qy      166 GAGAAAGTGTCTCGGTCTCCAAGATGGCGGCCGCTGGCCGTCTGGTCCGTCTGCTCCGGAG 225
          |||||||||||||||||||||||||||||||||||:::|||||||||||||||||
Db       3  GAGAAAGTGTCTCGGTCTCCAAGATGGCGGCCGCTGGCSDTCTGGTCCGTCTGCTCCGGAG 62

Qy      226 GCCGTGACGGCCAGACTCGTTGGTGTCTGTGGTTTCGTCTCAGTCACTACAGGACCCTGG 285
          |||||||||||||||||||||||||||||||||||
Db       63 GCCGTGACGGCCAGACTCGTTGGTGTCTGTGGTTTCGTCTCAGTCACTACAGGACCCTGG 122

Qy      286 GGGGCTGTTGCCACCTCCGCCGGGGGCGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTG 345
          |||||||||||||||||||||||||||||||||||
Db      123 GGGGCTGTTGCCACCTCCGCCGGGGGCGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTG 182

```

Qy 346 GGACAATATATTTGTAAAGATCCAAAAATAAATGACGCTACGCAAGAACCAGTTAACTGT 405  
 |||  
 Db 183 GGACAATATATTTGTAAAGATCCAAAAATAAATGACGCTACGCAAGAACCAGTTAACTGT 242

Qy 406 ACAAACTACACAGCTCATGTTTCCTGTTTCCAGCACCCAACATAACTTGTAAAGGATTCC 465  
 |||  
 Db 243 ACAAACTACACAGCTCATGTTTCCTGTTTCCAGCACCCAACATAACTTGTAAAGGATTCC 302

Qy 466 AGTGGCAATGAAACACATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGC 525  
 |||  
 Db 303 AGTGGCAATGAAACACATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGC 362

Qy 526 CGAAATGTAAATGGCTATTCTACAAAGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGG 585  
 |||  
 Db 363 CGAAATGTAAATGGCTATTCTACAAAGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGG 422

Qy 586 TTGGGAGCAGATCGATTT 603  
 |||  
 Db 423 TTGGGAGCAGATCGATTT 440

RESULT 13  
 AC102262/c

LOCUS AC102262 193660 bp DNA linear HTG 27-FEB-2003  
 DEFINITION Mus musculus clone RP24-216B4, WORKING DRAFT SEQUENCE, 9 unordered pieces.

ACCESSION AC102262  
 VERSION AC102262.3 GI:28570462  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 193660)  
 AUTHORS Birren,B., Nusbaum,C. and Lander,E.

TITLE Mus musculus, clone RP24-216B4  
 JOURNAL Unpublished

REFERENCE 2 (bases 1 to 193660)  
 AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
 Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,  
 Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,  
 Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,  
 Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,  
 Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,  
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
 Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,  
 Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,  
 Lamazares,R., Landers,T., Lehoczký,J., Levine,R., Liu,G.,  
 MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,  
 McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,  
 Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,  
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,  
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,  
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,  
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,  
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,

Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,  
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE Direct Submission  
 JOURNAL Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 REFERENCE 3 (bases 1 to 193660)  
 AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,  
 Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,  
 Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,  
 Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,  
 Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,  
 Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,  
 Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,  
 Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
 Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,  
 Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C.,  
 Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,  
 Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,  
 Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,  
 O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,  
 Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
 Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,  
 Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,  
 Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,  
 Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,  
 Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE Direct Submission  
 JOURNAL Submitted (27-FEB-2003) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 COMMENT On Feb 27, 2003 this sequence version replaced gi:22381123.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
 ----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: <http://www-seq.wi.mit.edu>  
 Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
 ----- Project Information  
 Center project name: L18275  
 Center clone name: 216\_B\_4  
 ----- Summary Statistics  
 Sequencing vector: Plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 191505 bases at least Q40  
 Consensus quality: 192178 bases at least Q30  
 Consensus quality: 192605 bases at least Q20  
 Insert size: 188000; agarose-fp  
 Insert size: 192860; sum-of-contigs  
 Quality coverage: 10.3 in Q20 bases; agarose-fp  
 Quality coverage: 10.0 in Q20 bases; sum-of-contigs  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 9 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

\* 1 16245: contig of 16245 bp in length  
 \* 16246 16345: gap of 100 bp  
 \* 16346 17545: contig of 1200 bp in length  
 \* 17546 17645: gap of 100 bp  
 \* 17646 21998: contig of 4353 bp in length  
 \* 21999 22098: gap of 100 bp  
 \* 22099 31133: contig of 9035 bp in length  
 \* 31134 31233: gap of 100 bp  
 \* 31234 41801: contig of 10568 bp in length  
 \* 41802 41901: gap of 100 bp  
 \* 41902 62223: contig of 20322 bp in length  
 \* 62224 62323: gap of 100 bp  
 \* 62324 91590: contig of 29267 bp in length  
 \* 91591 91690: gap of 100 bp  
 \* 91691 123816: contig of 32126 bp in length  
 \* 123817 123916: gap of 100 bp  
 \* 123917 193660: contig of 69744 bp in length.

FEATURES	Location/Qualifiers
source	1. .193660 /organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090" /clone="RP24-216B4"
misc_feature	1. .16245 /note="assembly_fragment clone_end:SP6 vector_side:left"
misc_feature	16346. .17545 /note="assembly_fragment"
misc_feature	17646. .21998 /note="assembly_fragment"
misc_feature	22099. .31133 /note="assembly_fragment"
misc_feature	31234. .41801 /note="assembly_fragment"
misc_feature	41902. .62223 /note="assembly_fragment"
misc_feature	62324. .91590 /note="assembly_fragment"
misc_feature	91691. .123816 /note="assembly_fragment"
misc_feature	123917. .193660 /note="assembly_fragment clone_end:T7 vector_side:right"

#### ORIGIN

Query Match 53.2%; Score 430.6; DB 2; Length 193660;  
 Best Local Similarity 81.7%; Pred. No. 5.2e-104;  
 Matches 535; Conservative 0; Mismatches 114; Indels 6; Gaps 3;

Qy 159 AAGTGGCGAGAAAGTGTCTGGTCTCCAAGATGGCGGCCGCTGGCCGTCTGGTCCGTCTGC 218  
 || ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 Db 138927 AAAAGGCTGGGAGGACGCAGTGCCCAACATGGCGGCTGCCTGGCCCGGGGCCGGGCTTC  
 138868

Qy 219 TCCGAGGCCGTGACGGCCAGACTCGTTGGTGTCTGTGGTTTCGTCTCAGTCACTACAGG 278  
 || | |||| | | | | | | | | | | | | | | | | | | | | | |  
 Db 138867 CCCAGCAGCCGGGCCTCCGGGCCTTCTCCGCACTCTGTGGCTCGTGATGGTCACCGAGGG  
 138808

Qy 279 ACCCTGGGGGGCTGTTGCCACCTCCGCCGGGGGCGAGGAGTCGCTTAAGTGCAGGACCT 338  
 || ||| |||||| | | | | | | | | | | | | | | | | | | | | | |  
 Db 138807 AACTGTGGGGCTGCTGCCTCTGGCGCTGTCGGGGGCGAGGAG---AGGTGTGAGGACCT  
 138751

Qy 339 CAAAGTGGGACAATATATTTGT--AAAGATCCAAAAATAAATGACGCTACGCAAGAACCA 396  
 || ||||||||| | | | | | | | | | | | | | | | | | | | | | | | |  
 Db 138750 CAGGGTGGGACAATATATTTGTAAAAAGAACCAAAAATAAATGATGCTACGCAAGAACCA  
 138691

Qy 397 GTTAACTGTACAACTACACAGCTCATGTTTCCTGTTTTCCAGCACCCAACATAACTTGT 456  
 |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 Db 138690 GTTAATTGTACAACTACACAGCTCATGTTCAATGTTTTCCAGCACCCAATAACTTAT  
 138631

Qy 457 -AAGGATTCCAGTGGCAATGAAACACATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCC 515  
 |||||| | | | | | | | | | | | | | | | | | | | | | | | | | |  
 Db 138630 AAAGGATTGAGTGGTAATGAAACACATTTTACTGGAAGTGAAGTCGGTTTTCTCAAGCC  
 138571

Qy 516 CATATCTTGCCGAAATGTAAATGGCTATTTCCTACAAAGTGGCAGTCGCATTGTCTCTTTT 575  
 |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 Db 138570 CATAACTTGCCGAAATGTGAATGGCTATTTCGTACAAAGTGGCAGTTGCATTATCTCTCTT  
 138511

Qy 576 TCTTGATGGTTGGGAGCAGATCGATTTTACCTTGGATACCCTGCTTTGGGTTTGTAAAA 635  
 | | |||| | | | | | | | | | | | | | | | | | | | | | | | | | |  
 Db 138510 TTTGGGATGGCTGGGAGCAGATCGATTTTACCTCGGATATCCTGCCTTACGCTTGTAAAA  
 138451

Qy 636 GTTTTGCAGTGTAGGGTTTTGTGGAATTGGGAGCCTAATTGATTTCACTTATTTCAT 695  
 |||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 Db 138450 TTTTGCACCGTAGGATTTTGCGGAATTGGGAGCCTAATTGATTTCACTTATTTCAT  
 138391

Qy 696 GCAGATTGTTGGACCTTCAGATGGAAGTAGTTACATTATAGATTACTATGGAACCAGACT 755  
 ||||||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 Db 138390 GCAGATTGTTGGACCTTCAGATGGAAGTAGTTACATTATAGACTATTATGGAACCAGGCT  
 138331

Qy 756 TACAAGACTGAGTATTACTAATGAAACATTTAGAAAAACGCAATTATATCCATAA 810  
 |||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 Db 138330 TACAAGACTCAGCATTACTAATGAAACATTTAGAAAAACCTAGCTGTACCCATAA 138276

LOCUS BD076249 455 bp DNA linear PAT 27-AUG-2002  
 DEFINITION 5' EST of tissue-nonspecific secretory protein.  
 ACCESSION BD076249  
 VERSION BD076249.1 GI:22621852  
 KEYWORDS JP 2001512011-A/197.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 455)  
 AUTHORS Edwards,J.B.D.M., Duclert,A. and Lacroix,B.  
 TITLE 5' EST of tissue-nonspecific secretory protein  
 JOURNAL Patent: JP 2001512011-A 197 21-AUG-2001;  
 GENSET  
 COMMENT OS Homo sapiens (human)  
 PN JP 2001512011-A/197  
 PD 21-AUG-2001  
 PF 31-JUL-1998 JP 2000505289  
 PR 01-AUG-1997 US 08/905135  
 PI JEAN BAPTISTE DUMAS MILNE EDWARDS,AYMERIC DUCLERT,BRUNO PI  
 LACROIX  
 PC C12N15/09,C12N15/09,C07K14/47,C12Q1/68,C12N15/00,C12N15/00 CC  
 blastn  
 CC identity 99  
 CC region 125. .358  
 CC id N47594  
 CC est  
 CC blastn  
 CC identity 100  
 CC region 49. .119  
 CC id N47594  
 CC est  
 CC blastn  
 CC identity 96  
 CC region 374. .438  
 CC id N47594  
 CC est  
 CC blastn  
 CC identity 96  
 CC region 113. .315  
 CC id AA143062  
 CC est  
 CC blastn  
 CC identity 97  
 CC region 43. .120  
 CC id AA143062  
 CC est  
 CC blastn  
 CC identity 98  
 CC region 304. .355  
 CC id AA143062  
 CC est  
 CC blastn  
 CC identity 95  
 CC region 371. .416  
 CC id AA143062  
 CC est



```

CC  blastn
CC  identity 97
CC  region 44. .317
CC  id HUM172D06B
CC  est
CC  blastn
CC  identity 97
CC  region 370. .416
CC  id HUM172D06B
CC  est
CC  blastn
CC  identity 97
CC  region 8. .46
CC  id HUM172D06B
CC  est
CC  blastn
CC  identity 97
CC  region 45. .359
CC  id HUM159G08B
CC  est
CC  blastn
CC  identity 97
CC  region 1. .47
CC  id HUM159G08B
CC  est
CC  blastn
CC  identity 98
CC  region 92. .316
CC  id N34957
CC  est
CC  blastn
CC  identity 100
CC  region 30. .97
CC  id N34957
CC  est
CC  Von Heijne matrix
CC  score 5.5
CC  seq LVGVLWVSVTTG/PW
CC  n=a, g, c or t
FH  Key          Location/Qualifiers
FT  misc_feature 141. .374
FT  misc_feature 65. .135
FT  misc_feature 388. .452
FT  misc_feature 131. .333
FT  misc_feature 60. .137
FT  misc_feature 323. .374
FT  misc_feature 388. .433
FT  misc_feature 60. .333
FT  misc_feature 388. .434
FT  misc_feature 23. .61
FT  misc_feature 60. .374
FT  misc_feature 15. .61
FT  misc_feature 131. .355
FT  misc_feature 68. .135
FT  sig_peptide   12. .104
FT  misc_feature 288
FT  misc_feature 375. .376

```

FT misc\_feature 385..387.  
 FEATURES Location/Qualifiers  
     source 1..455  
             /organism="Homo sapiens"  
             /mol\_type="genomic DNA"  
             /db\_xref="taxon:9606"

ORIGIN

Query Match 52.5%; Score 425.2; DB 6; Length 455;  
 Best Local Similarity 96.3%; Pred. No. 4.5e-103;  
 Matches 439; Conservative 3; Mismatches 12; Indels 2; Gaps 1;

```

Qy      177 GGTCTCCAAGATGGCGGCCGCCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGGC 236
          |||
Db      2   GGTCTCCAAGATGGCGGCCGCCTGGCCGTCTGGTCCGKCTGCTCCGGAGGCCGTGACGGC 61

Qy      237 CAGACTCGTTGGTGTCTCTGTGGTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGC 296
          |||
Db      62 CAGACTCGTTGGTGTCTCTGTGGTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGC 121

Qy      297 CACCTCCGCCCGGGGGCGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATAT 356
          |||
Db      122 CACCTCCGCCCGGGGGCGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATAT 181

Qy      357 TTGTAAAGATCCAAAATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACAC 416
          |||
Db      182 TTGTAAAGATCCAAAATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACAC 241

Qy      417 AGCTCATGTTTCCTGTTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGA 476
          |||
Db      242 AGCTCATGTTTCCTGTTTTCCAGCACCCAACATAACTTGTAAGGATNCCAGTGGCAATGA 301

Qy      477 AACACATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAA 536
          |||
Db      302 AACACATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAA 361

Qy      537 TGGCTATTTCCTACAAAGTGGCAGTCGCATTGTCTCTTTTTCTTGATGGTTGGGAGCAGA 596
          |||
Db      362 TGGCTATTTCCTAC--NNTKAGCAGTNNNWTGTCTCTTTTTCTTGATGGTTGGGAGCAGA 419

Qy      597 TCGATTTTACCTTGGATACCCTGCTTTGGGTTTGT 632
          |||
Db      420 TCGATTTTACCTTGGATACCCTGCTTTGGGTTTGT 455
  
```

RESULT 15

AX892343  
 LOCUS AX892343 487 bp DNA linear PAT 18-DEC-2003  
 DEFINITION Sequence 8206 from Patent EP1033401.  
 ACCESSION AX892343  
 VERSION AX892343.1 GI:40047227  
 KEYWORDS .  
 SOURCE Homo sapiens (human)  
     ORGANISM Homo sapiens  
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1

AUTHORS Dumas Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.  
TITLE Expressed sequence tags and encoded human proteins  
JOURNAL Patent: EP 1033401-A 8206 06-SEP-2000;  
Genset (FR)

FEATURES Location/Qualifiers  
source 1. .487  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 50.8%; Score 411.4; DB 6; Length 487;  
Best Local Similarity 91.4%; Pred. No. 2.3e-99;  
Matches 445; Conservative 9; Mismatches 10; Indels 23; Gaps 1;

```
Qy      164 GCGAGAAAGTGTCTGGTCTCCAAGATGGCGGCCCGCTGGCCGTCTGGTCCGTCTGCTCCGG 223
          |||
Db      1  GCGAGAAAGTGTCTGGTCTCCAAGATGGCGGCCCGCTGGACGTCTGGWCCGAMTGCACCGG 60

Qy      224 AGGCCGTGACGGCCAGACTCGTTGGTGTCTGTGGTTTCGTCTCAGTCACTACAGGACCCT 283
          |
Db      61 AAGCCGTGACGGCCAGAMTCGTTGGTGTCTGTGGTTTCGTMTCACTACTACAGGACCCT 120

Qy      284 GGGGGGCTGTTGCCACCTCCGCCGGGGGCGAGGAGTCGCTTAAGTGCGAGGACCTCAAAG 343
          |||
Db      121 GGGGGGCTGTTGCCACCTCCGCCGGGGGCRAGGAGTCGCTTAAGTGCGAGGACCTCAAAG 180

Qy      344 TGGGACAATATATT-----TGTAAGATCCAAAAATAAATGA 380
          |||
Db      181 TGRRACAATATCCTCTGTGGAGAACACCCCCCATGGAGGCGAGATCCAAAAATAAATGA 240

Qy      381 CGCTACGCAAGAACCAGTTAACTGTACAACTACACAGCTCATGTTTCCTGTTTCCAGC 440
          |||
Db      241 CGCTACGCAAGAACCAGTTAACTGTACAACTACACAGCTCATGTTTCCTGTTTCCAGC 300

Qy      441 ACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACACATTTTACTGGGAACGAAGT 500
          |||
Db      301 ACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACACATTTTACTGGGAACGAAGT 360

Qy      501 TGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCCTACAAAGTGGCAGT 560
          |||
Db      361 TGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCCTACAAAGTGGCAGT 420

Qy      561 CGCATTTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGGATACCCTGC 620
          |||
Db      421 CGCATTTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGGATACCCTGC 480

Qy      621 TTTGGGT 627
          |||
Db      481 TTTGGGT 487
```

Search completed: March 4, 2004, 08:34:00  
Job time : 3291 secs

OM nucleic - nucleic search, using sw model

Run on: March 4, 2004, 03:41:27 ; Search time 385 Seconds  
(without alignments)  
8937.767 Million cell updates/sec

Title: US-09-852-100B-1  
Perfect score: 810  
Sequence: 1 atgcatatttttaaagggtc.....aaacgcaattatatccataa 810

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002s:\*  
7: geneseqn2003as:\*  
8: geneseqn2003bs:\*  
9: geneseqn2003cs:\*  
10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

8							
Result	Query						
No.	Score	Match	Length	DB	ID	Description	
1	810	100.0	810	2	AAX05735	Aax05735	Human bet
2	810	100.0	810	3	AAZ52369	Aaz52369	Human bet
3	810	100.0	810	7	AAD51940	Aad51940	Human BBP
4	810	100.0	1246	7	AAD51979	Aad51979	Human BBP
5	602.8	74.4	970	2	AAX97705	Aax97705	Extended
6	499	61.6	508	3	AAA77946	Aaa77946	cDNA enco
7	499	61.6	508	3	AAA77958	Aaa77958	cDNA enco
8	499	61.6	508	4	AAI28684	Aai28684	Colon tum

	9	499	61.6	508	4	AAI28696	Aai28696	Colon tum
	10	499	61.6	508	7	ABZ32882	Abz32882	Human col
	11	499	61.6	508	7	ABZ32870	Abz32870	Human col
	12	442.2	54.6	1095	6	ABK52558	Abk52558	cDNA enco
	13	436.8	53.9	440	2	AAX41191	Aax41191	Human sec
	14	425.2	52.5	455	2	AAX41259	Aax41259	Human sec
	15	411.4	50.8	487	3	AAC04131	Aac04131	Human sec
	16	337.4	41.7	431	8	ACH19706	Ach19706	Human adu
	17	322	39.8	323	7	ACD92727	Acd92727	Human col
	18	320.4	39.6	323	7	ACD92728	Acd92728	Human col
c	19	182.4	22.5	561	6	ABQ57089	Abq57089	Human col
	20	153.2	18.9	515	9	ADB56631	Adb56631	Toxicity-
	21	115.4	14.2	433	5	ABV17809	Abv17809	Human pro
	22	115.4	14.2	487	5	ABV47601	Abv47601	Human pro
	23	113	14.0	292	7	AAD51978	Aad51978	Human BBP
	24	99.8	12.3	298	2	AAX85735	Aax85735	Novel cDN
c	25	64.6	8.0	2771	4	ABL16838	Abl16838	Drosophil
c	26	64.6	8.0	3642	4	ABL15742	Abl15742	Drosophil
	27	64.4	8.0	706	4	ABL16839	Abl16839	Drosophil
	28	51	6.3	1369	2	AAX85024	Aax85024	Human sec
	29	51	6.3	1369	7	ADA56147	Ada56147	Gene enco
	30	51	6.3	1369	7	ACD18950	Acd18950	Novel hum
	31	51	6.3	1369	7	ACC50529	Acc50529	Human sec
	32	51	6.3	1369	7	ABZ71291	Abz71291	Secreted
	33	51	6.3	1369	8	ADB91205	Adb91205	Human sec
	34	51	6.3	1369	9	ADC73589	Adc73589	Human sec
	35	50.4	6.2	690	3	AAA64413	Aaa64413	Open read
	36	50.4	6.2	854	3	AAA64412	Aaa64412	DNA encod
c	37	50.2	6.2	439	6	ABK62922	Abk62922	Rat seque
c	38	50.2	6.2	439	9	ADB56953	Adb56953	Toxicity-
	39	49.8	6.1	741	3	AAA64409	Aaa64409	Open read
	40	49.8	6.1	746	3	AAZ52371	Aaz52371	Human bet
	41	49.8	6.1	1406	7	ACC51100	Acc51100	Human Amy
	42	49.8	6.1	1455	4	AAF80523	Aaf80523	Receptor
	43	49.8	6.1	1473	3	AAA64408	Aaa64408	DNA encod
	44	49.8	6.1	1473	3	AAA64425	Aaa64425	DNA encod
	45	49.8	6.1	1473	3	AAA64424	Aaa64424	DNA encod

#### ALIGNMENTS

##### RESULT 1

AAX05735

ID AAX05735 standard; mRNA; 810 BP.

XX

AC AAX05735;

XX

DT 27-APR-1999 (first entry)

XX

DE Human beta-amyloid peptide-binding protein (BBP) encoding mRNA.

XX

KW Beta-amyloid peptide binding protein; BBP; beta-amyloid protein; BAP;

KW human; Alzheimer's disease; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers  
 FT CDS 1. .810  
 FT /\*tag= a  
 FT /product= "BBP"  
 XX  
 PN WO9846636-A2.  
 XX  
 PD 22-OCT-1998.  
 XX  
 PF 14-APR-1998; 98WO-US007462.  
 XX  
 PR 16-APR-1997; 97US-0064583P.  
 XX  
 PA (AMHP ) AMERICAN HOME PROD CORP.  
 XX  
 PI Ozenberger BA, Kajkowski EM, Jacobsen JS, Bard JA, Walker SG;  
 XX  
 DR WPI; 1999-080736/07.  
 DR P-PSDB; AAW94291.  
 XX  
 PT Polynucleotide encoding beta-amyloid peptide binding protein - used to  
 PT identify inhibitors of beta-amyloid peptide for treating Alzheimer's  
 PT disease.  
 XX  
 PS Claim 1; Page 43-44; 59pp; English.  
 XX  
 CC This represents a nucleotide sequence encoding a beta-amyloid peptide  
 CC binding protein (BBP). The polynucleotide comprising the entire BBP  
 CC nucleotide sequence of clone BBP1-fl is deposited under the accession  
 CC number ATCC 98617. The polynucleotide comprising a fragment of BBP  
 CC (nucleotides 202-807 of the full length BBP) of clone pEK196 is deposited  
 CC as ATCC 98399. Host cells transformed with a vector comprising the BBP  
 CC nucleic acid are used for the recombinant production of the protein. The  
 CC protein can be used in a method for diagnosing a disease characterised by  
 CC aberrant expression of human beta-amyloid protein (BAP). The protein can  
 CC also be used in a method for screening for compounds which regulate  
 CC expression of a BAP binding protein. The proteins, antibodies and  
 CC identified compounds can be used in the treatment or prevention of  
 CC Alzheimer's disease  
 XX  
 SQ Sequence 810 BP; 204 A; 183 C; 202 G; 221 T; 0 U; 0 Other;

Query Match 100.0%; Score 810; DB 2; Length 810;  
 Best Local Similarity 100.0%; Pred. No. 2e-233;  
 Matches 810; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGCATATTTTAAAAGGGTCTCCCAATGTGATTCCACGGGCTCAGGGGCAGAAGAACACG	60
Db	1	ATGCATATTTTAAAAGGGTCTCCCAATGTGATTCCACGGGCTCAGGGGCAGAAGAACACG	60
Qy	61	CGAAGAGACGGAACCTGGCCTCTATCCTATGCGAGGTCCCTTTAAGAACCTCGCCCTGTTG	120
Db	61	CGAAGAGACGGAACCTGGCCTCTATCCTATGCGAGGTCCCTTTAAGAACCTCGCCCTGTTG	120
Qy	121	CCCTTCTCCCTCCCGCTCCTGGGCGGAGGCGGAAGCGGAAGTGGCGAGAAAGTGTCGGTC	180
Db	121	CCCTTCTCCCTCCCGCTCCTGGGCGGAGGCGGAAGCGGAAGTGGCGAGAAAGTGTCGGTC	180

Qy	181	TCCAAGATGGCGGCCCGCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGGCCAGA	240
Db	181	TCCAAGATGGCGGCCCGCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGGCCAGA	240
Qy	241	CTCGTTGGTGTCTGTGGTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGCCACC	300
Db	241	CTCGTTGGTGTCTGTGGTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGCCACC	300
Qy	301	TCCGCCGGGGGCGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATATTTGT	360
Db	301	TCCGCCGGGGGCGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATATTTGT	360
Qy	361	AAAGATCCAAAATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACACAGCT	420
Db	361	AAAGATCCAAAATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACACAGCT	420
Qy	421	CATGTTTCCTGTTTTCCAGCACCCAACATAAAGTGTAAAGGATTCCAGTGGCAATGAAACA	480
Db	421	CATGTTTCCTGTTTTCCAGCACCCAACATAAAGTGTAAAGGATTCCAGTGGCAATGAAACA	480
Qy	481	CATTTTACTGGGAACGAAGTTGGTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGC	540
Db	481	CATTTTACTGGGAACGAAGTTGGTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGC	540
Qy	541	TATTCCTACAAAGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGA	600
Db	541	TATTCCTACAAAGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGA	600
Qy	601	TTTTACCTTGGATACCCTGCTTTGGGTTTGTAAAGTTTTGCACTGTAGGGTTTTGTGGA	660
Db	601	TTTTACCTTGGATACCCTGCTTTGGGTTTGTAAAGTTTTGCACTGTAGGGTTTTGTGGA	660
Qy	661	ATTGGGAGCCTAATTGATTTCAATTCCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGA	720
Db	661	ATTGGGAGCCTAATTGATTTCAATTCCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGA	720
Qy	721	AGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAA	780
Db	721	AGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAA	780
Qy	781	ACATTTAGAAAAACGCAATTATATCCATAA	810
Db	781	ACATTTAGAAAAACGCAATTATATCCATAA	810

---

RESULT 2

AAZ52369

ID AAZ52369 standard; cDNA; 810 BP.

XX

AC AAZ52369;

XX

DT 24-JUL-2000 (first entry)

XX

DE Human beta-amyloid peptide (BAP) binding protein, BBP1 encoding cDNA.

XX

KW Beta-amyloid peptide binding protein; BBP; BAP; tumour; suppressor;

KW G-protein coupled receptor; GPCR; integral membrane protein; antigen;  
 KW neuronal cell; nonhuman primate; NHP; G-protein signalling pathway;  
 KW apoptosis; immunogen; therapeutic; treatment; prevention; diagnostic; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1. .810  
 FT /\*tag= a  
 FT /product= "Human BBP1 protein"  
 FT /note= "Member of G-protein coupled receptor superfamily"  
 XX  
 PN WO200022125-A2.  
 XX  
 PD 20-APR-2000.  
 XX  
 PF 13-OCT-1999; 99WO-US021621.  
 XX  
 PR 13-OCT-1998; 98US-0104104P.  
 XX  
 PA (AMHP ) AMERICAN HOME PROD CORP.  
 XX  
 PI Ozenberger BA, Kajkowski EM, Lo CF;  
 XX  
 DR WPI; 2000-317982/27.  
 DR P-PSDB; AAY70759.  
 XX  
 PT Novel G-protein-coupled receptor-like proteins and polynucleotides useful  
 PT for regulating apoptosis, comprises integral membrane protein traversing  
 PT the membrane twice.  
 XX  
 PS Example 1; Page 60-61; 68pp; English.  
 XX  
 CC The present sequence is the cDNA encoding beta-amyloid peptide (BAP)  
 CC binding protein-1 (BBP1). It is an integral membrane protein, that  
 CC traverse the membrane twice. It is related to G protein-coupled receptor  
 CC (GPCR) protein superfamily. It interacts with G-alpha proteins and  
 CC regulates the activity of G-protein signalling pathways. BBP genes are  
 CC widely expressed in neuronal cells of nonhuman primate (NHP) brain and  
 CC overexpressed in some tumours. It functions as a suppressor of apoptosis  
 CC induction. BBP proteins are used as immunogens to raise antibodies,  
 CC useful as therapeutics and as antigens in solid phase assays. They are  
 CC also useful as reagents to identify molecules which effect the  
 CC interaction of BBP and a cloned protein, that are useful in the treatment  
 CC or prevention of diseases associated with apoptosis. The polynucleotides  
 CC are useful for diagnostics. Note: In claim 5, the patent claims an amino  
 CC acid sequence from figure 2. However, figure 2 does not contain any  
 CC sequence. It is inferred from the disclosure that the figure 2 sequence  
 CC refers to BBP1 protein, encoded by this polynucleotide sequence  
 XX  
 SQ Sequence 810 BP; 204 A; 183 C; 202 G; 221 T; 0 U; 0 Other;

Query Match 100.0%; Score 810; DB 3; Length 810;  
 Best Local Similarity 100.0%; Pred. No. 2e-233;  
 Matches 810; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCATATTTTAAAAGGGTCTCCCAATGTGATTCCACGGGCTCACGGGCAGAAGAACACG 60



Db	1	ATGCATATTTTAAAAGGGTCTCCCAATGTGATTCCACGGGCTCACGGGCAGAAGAACACG	60
Qy	61	CGAAGAGACGGAACGGCCTCTATCCTATGCGAGGTCCCTTTAAGAACCTCGCCCTGTTG	120
Db	61	CGAAGAGACGGAACGGCCTCTATCCTATGCGAGGTCCCTTTAAGAACCTCGCCCTGTTG	120
Qy	121	CCCTTCTCCCTCCCGCTCCTGGGCGGAGGCGGAAGCGGAAGTGGCGAGAAAGTGTCTGGTC	180
Db	121	CCCTTCTCCCTCCCGCTCCTGGGCGGAGGCGGAAGCGGAAGTGGCGAGAAAGTGTCTGGTC	180
Qy	181	TCCAAGATGGCGGCCCGCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGGCCAGA	240
Db	181	TCCAAGATGGCGGCCCGCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGGCCAGA	240
Qy	241	CTCGTTGGTGTCTGTGGTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGCCACC	300
Db	241	CTCGTTGGTGTCTGTGGTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGCCACC	300
Qy	301	TCCGCCGGGGGCGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGT	360
Db	301	TCCGCCGGGGGCGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGT	360
Qy	361	AAAGATCCAAAAATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACACAGCT	420
Db	361	AAAGATCCAAAAATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACACAGCT	420
Qy	421	CATGTTTCCTGTTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACA	480
Db	421	CATGTTTCCTGTTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACA	480
Qy	481	CATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGC	540
Db	481	CATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGC	540
Qy	541	TATTCCTACAAAGTGGCAGTCGCATTGTCTCTTTTTCTTGATGGTTGGGAGCAGATCGA	600
Db	541	TATTCCTACAAAGTGGCAGTCGCATTGTCTCTTTTTCTTGATGGTTGGGAGCAGATCGA	600
Qy	601	TTTTACCTTGGATACCCTGCTTTGGGTTTGTTAAAGTTTTGCACTGTAGGGTTTTGTGGA	660
Db	601	TTTTACCTTGGATACCCTGCTTTGGGTTTGTTAAAGTTTTGCACTGTAGGGTTTTGTGGA	660
Qy	661	ATTGGGAGCCTAATTGATTTCACTTATTTCAATGCAGATTGTTGGACCTTCAGATGGA	720
Db	661	ATTGGGAGCCTAATTGATTTCACTTATTTCAATGCAGATTGTTGGACCTTCAGATGGA	720
Qy	721	AGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAA	780
Db	721	AGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAA	780
Qy	781	ACATTTAGAAAAACGCAATTATATCCATAA	810
Db	781	ACATTTAGAAAAACGCAATTATATCCATAA	810

AAD51940

ID AAD51940 standard; cDNA; 810 BP.

XX

AC AAD51940;

XX

DT 02-MAY-2003 (first entry)

XX

DE Human BBP-1 cDNA.

XX

KW Human; beta-amyloid peptide-binding protein; BAP; Abeta; betaAP; BBP;

KW Alzheimer's disease; AD; transgenic; transgenic animal; gene therapy;

KW neuroprotective; nootropic; gene; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 1. .810

FT /\*tag= a

FT /product= "Human BBP-1"

XX

PN WO200290499-A2.

XX

PD 14-NOV-2002.

XX

PF 06-MAY-2002; 2002WO-US014223.

XX

PR 09-MAY-2001; 2001US-00852100.

XX

PA (AMHP ) WYETH.

XX

PI Ozenberger BA, Bard JA, Kajkowski EM, Jacobsen JS, Walker SG;

PI Sofia HJ, Howland DS;

XX

DR WPI; 2003-120537/11.

DR P-PSDB; AAE33877.

XX

PT New human beta-amyloid peptide-binding protein, useful for diagnosing

PT and/or treating diseases associated with aberrant expression of beta-

PT amyloid peptide, e.g. Alzheimer's disease.

XX

PS Claim 1; Page 82-84; 85pp; English.

XX

CC The present invention relates to novel human beta-amyloid peptide (BAP;

CC Abeta, betaAP)-binding (BBP) proteins and polynucleotides encoding such

CC proteins. BBP sequences are useful to diagnose and/or treat diseases

CC associated with aberrant expression of human BAP such as Alzheimer's

CC disease (AD). They are used to generate transgenic animals. Sequences of

CC the invention are also used in gene therapy. The present sequence is

CC human BBP-1 cDNA

XX

SQ Sequence 810 BP; 204 A; 183 C; 202 G; 221 T; 0 U; 0 Other;

Query Match 100.0%; Score 810; DB 7; Length 810;

Best Local Similarity 100.0%; Pred. No. 2e-233;

Matches 810; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

1 ATGCATATTTTAAAGGGTCTCCCAATGTGATTCCACGGGCTCACGGGCAGAAGAACACG 60

Db	1		ATGCATATTTTAAAGGGTCTCCCAATGTGATTCCACGGGCTCACGGGCAGAAGAACACG	60
Qy	61		CGAAGAGACGGAACCTGGCCTCTATCCTATGCGAGGTCCCTTTAAGAACCTCGCCCTGTTG	120
Db	61		CGAAGAGACGGAACCTGGCCTCTATCCTATGCGAGGTCCCTTTAAGAACCTCGCCCTGTTG	120
Qy	121		CCCTTCTCCCTCCCGCTCCTGGGCGGAGGCGGAAGCGGAAGTGGCGAGAAAGTGTCCGGTC	180
Db	121		CCCTTCTCCCTCCCGCTCCTGGGCGGAGGCGGAAGCGGAAGTGGCGAGAAAGTGTCCGGTC	180
Qy	181		TCCAAGATGGCGGCCCGCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGGCCAGA	240
Db	181		TCCAAGATGGCGGCCCGCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGGCCAGA	240
Qy	241		CTCGTTGGTGTCTGTGGTTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGCCACC	300
Db	241		CTCGTTGGTGTCTGTGGTTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGCCACC	300
Qy	301		TCCGCCGGGGCGGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGT	360
Db	301		TCCGCCGGGGCGGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGT	360
Qy	361		AAAGATCCAAAAATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACACAGCT	420
Db	361		AAAGATCCAAAAATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACACAGCT	420
Qy	421		CATGTTTCCTGTTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACA	480
Db	421		CATGTTTCCTGTTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACA	480
Qy	481		CATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGC	540
Db	481		CATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGC	540
Qy	541		TATTCTACAAAGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGA	600
Db	541		TATTCTACAAAGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGA	600
Qy	601		TTTTACCTTGGATACCCTGCTTTGGGTTTGTTAAAGTTTTGCACTGTAGGGTTTTGTGGA	660
Db	601		TTTTACCTTGGATACCCTGCTTTGGGTTTGTTAAAGTTTTGCACTGTAGGGTTTTGTGGA	660
Qy	661		ATTGGGAGCCTAATTGATTTCAATCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGA	720
Db	661		ATTGGGAGCCTAATTGATTTCAATCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGA	720
Qy	721		AGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAA	780
Db	721		AGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAA	780
Qy	781		ACATTTAGAAAAACGCAATTATATCCATAA	810
Db	781		ACATTTAGAAAAACGCAATTATATCCATAA	810

```

AAD51979
ID  AAD51979 standard; DNA; 1246 BP.
XX
AC  AAD51979;
XX
DT  02-MAY-2003 (first entry)
XX
DE  Human BBP-1 genomic DNA.
XX
KW  Human; beta-amyloid peptide-binding protein; BAP; Abeta; betaAP; BBP;
KW  Alzheimer's disease; AD; transgenic; transgenic animal; gene therapy;
KW  neuroprotective; nootropic; ds.
XX
OS  Homo sapiens.
XX
PN  WO200290499-A2.
XX
PD  14-NOV-2002.
XX
PF  06-MAY-2002; 2002WO-US014223.
XX
PR  09-MAY-2001; 2001US-00852100.
XX
PA  (AMHP ) WYETH.
XX
PI  Ozenberger BA, Bard JA, Kajkowski EM, Jacobsen JS, Walker SG;
PI  Sofia HJ, Howland DS;
XX
DR  WPI; 2003-120537/11.
XX
PT  New human beta-amyloid peptide-binding protein, useful for diagnosing
PT  and/or treating diseases associated with aberrant expression of beta-
PT  amyloid peptide, e.g. Alzheimer's disease.
XX
PS  Disclosure; Fig 11; 85pp; English.
XX
CC  The present invention relates to novel human beta-amyloid peptide (BAP;
CC  Abeta, betaAP)-binding (BBP) proteins and polynucleotides encoding such
CC  proteins. BBP sequences are useful to diagnose and/or treat diseases
CC  associated with aberrant expression of human BAP such as Alzheimer's
CC  disease (AD). They are used to generate transgenic animals. Sequences of
CC  the invention are also used in gene therapy. The present sequence is
CC  human BBP-1 genomic DNA
XX
SQ  Sequence 1246 BP; 318 A; 255 C; 283 G; 390 T; 0 U; 0 Other;

```

Qy	121	CCCTTCTCCCTCCCGCTCCTGGGCGGAGGCGGAAGCGGAAGTGGCGAGAAAGTGTGGTC	180
Db	238	CCCTTCTCCCTCCCGCTCCTGGGCGGAGGCGGAAGCGGAAGTGGCGAGAAAGTGTGGTC	297
Qy	181	TCCAAGATGGCGGCCCGCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGGCCAGA	240
Db	298	TCCAAGATGGCGGCCCGCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGGCCAGA	357
Qy	241	CTCGTTGGTGTCTGTGGTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGCCACC	300
Db	358	CTCGTTGGTGTCTGTGGTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGCCACC	417
Qy	301	TCCGCCGGGGGCGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATATTTGT	360
Db	418	TCCGCCGGGGGCGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATATTTGT	477
Qy	361	AAAGATCCAAAAATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCT	420
Db	478	AAAGATCCAAAAATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCT	537
Qy	421	CATGTTTCCTGTTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACA	480
Db	538	CATGTTTCCTGTTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACA	597
Qy	481	CATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGC	540
Db	598	CATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGC	657
Qy	541	TATTCCTACAAAGTGGCAGTCGCATTGTCTCTTTTTCTTGATGGTTGGGAGCAGATCGA	600
Db	658	TATTCCTACAAAGTGGCAGTCGCATTGTCTCTTTTTCTTGATGGTTGGGAGCAGATCGA	717
Qy	601	TTTTACCTTGATACCCCTGCTTTGGGTTTGTTAAAGTTTGCAGTGTAGGGTTTGTGGA	660
Db	718	TTTTACCTTGATACCCCTGCTTTGGGTTTGTTAAAGTTTGCAGTGTAGGGTTTGTGGA	777
Qy	661	ATTGGGAGCCTAATTGATTTTCAATCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGA	720
Db	778	ATTGGGAGCCTAATTGATTTTCAATCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGA	837
Qy	721	AGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAA	780
Db	838	AGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAA	897
Qy	781	ACATTTAGAAAAACGCAATTATATCCATAA	810
Db	898	ACATTTAGAAAAACGCAATTATATCCATAA	927

RESULT 5

AAX97705

ID AAX97705 standard; DNA; 970 BP.

XX

AC AAX97705;

XX

DT 13-SEP-1999 (first entry)

XX  
 DE Extended human secreted protein coding sequence, SEQ ID NO. 270.  
 XX  
 KW Secreted protein; human; cytokine; cellular proliferation; cell movement;  
 KW cellular differentiation; immune system regulator; anti-inflammatory;  
 KW haematopoiesis regulator; tissue growth regulator; tumour inhibitor;  
 KW reproductive hormone regulator; chemotaxis; chemokinesis; gene therapy;  
 KW genetic disease; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9931236-A2.  
 XX  
 PD 24-JUN-1999.  
 XX  
 PF 17-DEC-1998; 98WO-IB002122.  
 XX  
 PR 17-DEC-1997; 97US-0069957P.  
 PR 09-FEB-1998; 98US-0074121P.  
 PR 13-APR-1998; 98US-0081563P.  
 PR 10-AUG-1998; 98US-0096116P.  
 XX  
 PA (GEST ) GENSET.  
 XX  
 PI Bougueleret L, Duclert A, Dumas Milne Edwards J;  
 XX  
 DR WPI; 1999-385906/32.  
 DR P-PSDB; AAY36021.  
 XX  
 PT New isolated human secreted proteins.  
 XX  
 PS Claim 1; Page 346-347; 516pp; English.  
 XX  
 CC This sequence represents an extended human secreted protein coding  
 CC sequence of the invention. The secreted proteins can be used in treating  
 CC or controlling a variety of human conditions. The secreted proteins may  
 CC act as cytokines or may affect cellular proliferation or differentiation  
 CC or may act as immune system regulators, haematopoiesis regulators, tissue  
 CC growth regulators, regulators of reproductive hormones or cell movement  
 CC or have chemotactic/chemokinetic, receptor/ligand, anti-inflammatory or  
 CC tumour inhibition activity. The DNAs can be used in forensic procedures  
 CC to identify individuals or in diagnostic procedures to identify  
 CC individuals having genetic diseases resulting from abnormal expression of  
 CC the genes corresponding to the extended cDNAs. They are also useful for  
 CC constructing a high resolution map of the human chromosomes. They can  
 CC also be used for gene therapy to control or treat genetic diseases  
 XX  
 SQ Sequence 970 BP; 267 A; 173 C; 199 G; 323 T; 0 U; 8 Other;  
  
 Query Match 74.4%; Score 602.8; DB 2; Length 970;  
 Best Local Similarity 98.4%; Pred. No. 6.7e-171;  
 Matches 624; Conservative 5; Mismatches 3; Indels 2; Gaps 2;  
  
 Qy 177 GGTCTCCAAGATGGCGGCCCGCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGGC 236  
 ||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||  
 Db 2 GGTCTCCAAGATGGCGGCCCGCTGGCCGTCTGGTCCGKCTGCTCCGGAGGCCGTGACGGC 61

Qy	237	CAGACTCGTTGGTGTCTGTGGTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGC	296
Db	62	CAGACTCGTTGGTGTCTGTGGTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGC	121
Qy	297	CACCTCCGCCGGGGGCGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATAT	356
Db	122	CACCTCCGCCGGGGGCGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATAT	181
Qy	357	TTGTAAAGATCCAAAAATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACAC	416
Db	182	TTGTAAAGATCCAAAAATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACAC	241
Qy	417	AGCTCATGTTTCCTGTTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGA	476
Db	242	AGCTCATGTTTCCTGTTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGA	301
Qy	477	AACACATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAA	536
Db	302	AACACATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAA	361
Qy	537	TGGCTATTCTACAAAGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGA	596
Db	362	TGGCTATTCTACAAAGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGA	419
Qy	597	TCGATTTTACCTTGGATACCCTGCCTTTGGGTTTGTAAAGTTTTGCACTGTAGGGTTTTG	656
		: : :	
Db	420	TCGATTTTACCTTGGATACCCTGCCTTTGGGTTTGTAAABTTTYGCACTGTAGGGTTTTKG	479
Qy	657	TGGAATTGGGAGCCTAATTGATTTCACTTATTTCAATGCAGATTGTTGGACCTTCAGA	716
		:	
Db	480	TGGAATTGGGAGCCTAATTGATTTCACTTATTTCAATGCAGATTGTTGGACCTTCAGA	539
Qy	717	TGGAAGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAA	776
Db	540	TGGAAGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAA	599
Qy	777	TGAAACATTTAGAAAAACGCAATTATATCCATAA	810
Db	600	TGAAACATTTAGAAAAACGCAATTATATCCATAA	633

# RESULT 6

AAA77946

ID AAA77946 standard; cDNA; 508 BP.

XX

AC AAA77946;

XX

DT 14-NOV-2000 (first entry)

XX

DE cDNA encoding human colon tumour polypeptide, SEQ ID NO:233.

XX

KW Human colon tumour polypeptide; tumour antigen; cancer; vaccine;

KW immunotherapy; diagnosis; progression; ss.

XX

OS Homo sapiens.

XX

PN WO200037643-A2.

```

XX PD 29-JUN-2000.
XX
XX PF 23-DEC-1999; 99WO-US030909.
XX
XX PR 23-DEC-1998; 98US-00221298.
XX PR 02-JUL-1999; 99US-00347496.
XX PR 22-SEP-1999; 99US-00401064.
XX PR 19-NOV-1999; 99US-00444242.
XX PR 02-DEC-1999; 99US-00454150.
XX
XX PA (CORI-) CORIXA CORP.
XX
XX PI Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk J;
XX PI Wang T, Yuqiu J;
XX
XX DR WPI; 2000-442671/38.
XX
XX PT New colon tumor polypeptides used to inhibit the development of cancer,
XX PT especially colon cancer, and for diagnosing and monitoring the
XX PT progression of the cancer.
XX
XX PS Claim 1; Page 158-159; 229pp; English.
XX
XX CC Sequences AAA77722-A78199 represent 478 cDNAs encoding proteins or
XX CC portions of proteins which are associated with human colon tumours. The
XX CC invention also specifically discloses 8 human colon tumour proteins
XX CC (AAB11897-B11904). The nucleic acids, the polypeptides they encode, and
XX CC antigen presenting cells (APCs, preferably dendritic cells) expressing
XX CC such polypeptides may be used in vaccines that target tumour cells,
XX CC especially colon tumour cells, thereby inhibiting the development of
XX CC cancer. T-cells specific for the polypeptide expressed by the APC are
XX CC used to remove tumour cells from biological samples, especially blood or
XX CC fractions thereof. The sample or the isolated T-cells specific for the
XX CC polypeptide can then be used to inhibit cancer development. CD4+ and/or
XX CC CD8+ T-cells from a patient may be incubated with a polypeptide or
XX CC nucleic acid of the invention, or an APC expressing such a polypeptide,
XX CC to cause the proliferation of specific T-cells. The T-cells can be cloned
XX CC and then administered back to the patient to inhibit cancer development.
XX CC Nucleic acids encoding the polypeptides and antibodies against the
XX CC polypeptides may be used to determine the expression level of a tumour
XX CC protein of the invention, and therefore to determine whether cancer cells
XX CC are present. Such diagnostic methods may also be used to monitor the
XX CC progression of a cancer by repeating the processes at time intervals, and
XX CC comparing the current result to previous results. The present sequence
XX CC represents a cDNA encoding a human colon tumour polypeptide
XX
XX SQ Sequence 508 BP; 153 A; 89 C; 103 G; 163 T; 0 U; 0 Other;
XX
XX Query Match 61.6%; Score 499; DB 3; Length 508;
XX Best Local Similarity 100.0%; Pred. No. 9.4e-140;
XX Matches 499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 312 CGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 371
XX |
XX Db 1 CGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 60

```



Qy	372	AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCTCATGTTTCCTG	431
Db	61	AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCTCATGTTTCCTG	120
Qy	432	TTTTCCAGCACCCAACATAAAGGATTCCAGTGGCAATGAAACACATTTTACTGG	491
Db	121	TTTTCCAGCACCCAACATAAAGGATTCCAGTGGCAATGAAACACATTTTACTGG	180
Qy	492	GAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCTTACAA	551
Db	181	GAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCTTACAA	240
Qy	552	AGTGGCAGTCGCATTGTCTCTTTTTCTTGATGGTTGGGAGCAGATCGATTTTACCTTGG	611
Db	241	AGTGGCAGTCGCATTGTCTCTTTTTCTTGATGGTTGGGAGCAGATCGATTTTACCTTGG	300
Qy	612	ATACCCTGCTTTGGGTTTGTAAAGTTTTGCACTGTAGGGTTTTGTGGAATTGGGAGCCT	671
Db	301	ATACCCTGCTTTGGGTTTGTAAAGTTTTGCACTGTAGGGTTTTGTGGAATTGGGAGCCT	360
Qy	672	AATTGATTTCATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT	731
Db	361	AATTGATTTCATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT	420
Qy	732	TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA	791
Db	421	TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA	480
Qy	792	AACGCAATTATATCCATAA	810
Db	481	AACGCAATTATATCCATAA	499

# RESULT 7

AAA77958

ID AAA77958 standard; cDNA; 508 BP.

XX

AC AAA77958;

XX

DT 14-NOV-2000 (first entry)

XX

DE cDNA encoding human colon tumour polypeptide, SEQ ID NO:245.

XX

KW Human colon tumour polypeptide; tumour antigen; cancer; vaccine;

KW immunotherapy; diagnosis; progression; ss.

XX

OS Homo sapiens.

XX

PN WO200037643-A2.

XX

PD 29-JUN-2000.

XX

PF 23-DEC-1999; 99WO-US030909.

XX

PR 23-DEC-1998; 98US-00221298.

PR 02-JUL-1999; 99US-00347496.

PR 22-SEP-1999; 99US-00401064.

PR 02-DEC-1999; 99US-00454150.

XX

PA (CORI-) CORIXA CORP.

XX

PI Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk J;

PI Wang T, Yuqiu J;

XX

DR WPI: 2000-442671/38.

XX

PT New colon tumor polypeptides used to inhibit the development of cancer,

PT especially colon cancer, and for diagnosing and monitoring the

PT progression of the cancer.

XX

PS Claim 1; Page 162; 229pp; English.

XX

Sequences AAA77722-A78199 represent 478 cDNAs encoding proteins or portions of proteins which are associated with human colon tumours. The invention also specifically discloses 8 human colon tumour proteins (AAB11897-B11904). The nucleic acids, the polypeptides they encode, and antigen presenting cells (APCs, preferably dendritic cells) expressing such polypeptides may be used in vaccines that target tumour cells, especially colon tumour cells, thereby inhibiting the development of cancer. T-cells specific for the polypeptide expressed by the APC are used to remove tumour cells from biological samples, especially blood or fractions thereof. The sample or the isolated T-cells specific for the polypeptide can then be used to inhibit cancer development. CD4+ and/or CD8+ T-cells from a patient may be incubated with a polypeptide or nucleic acid of the invention, or an APC expressing such a polypeptide, to cause the proliferation of specific T-cells. The T-cells can be cloned and then administered back to the patient to inhibit cancer development. Nucleic acids encoding the polypeptides and antibodies against the polypeptides may be used to determine the expression level of a tumour protein of the invention, and therefore to determine whether cancer cells are present. Such diagnostic methods may also be used to monitor the progression of a cancer by repeating the processes at time intervals, and comparing the current result to previous results. The present sequence represents a cDNA encoding a human colon tumour polypeptide

XX

Sequence 508 BP; 153 A; 89 C; 103 G; 163 T; 0 U; 0 Other;

Query Match 61.6%; Score 499; DB 3; Length 508;

Best Local Similarity 100.0%; Pred. No. 9.4e-140;

Matches 499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 312 CGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 371

Db 1 CGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 60

Ov 372 AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCTCATGTTTCCTG 431

Db 61 AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACACAGCTCATGTTTCCTG 120

Ov 432 TTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 491

Db 121 TTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 180

Qy 492 GAACGAAGTTGGTTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCTTACAA 551  
 |||  
 Db 181 GAACGAAGTTGGTTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCTTACAA 240  
 Qy 552 AGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTGGGAGCAGATCGATTTTACCTTGG 611  
 |||  
 Db 241 AGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTGGGAGCAGATCGATTTTACCTTGG 300  
 Qy 612 ATACCCTGCTTTGGGTTTGTAAAGTTTTGCACTGTAGGGTTTGTGGAATTGGGAGCCT 671  
 |||  
 Db 301 ATACCCTGCTTTGGGTTTGTAAAGTTTTGCACTGTAGGGTTTGTGGAATTGGGAGCCT 360  
 Qy 672 AATTGATTTCAATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 731  
 |||  
 Db 361 AATTGATTTCAATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 420  
 Qy 732 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 791  
 |||  
 Db 421 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 480  
 Qy 792 AACGCAATTATATCCATAA 810  
 |||  
 Db 481 AACGCAATTATATCCATAA 499

RESULT 8

AAI28684

ID AAI28684 standard; cDNA; 508 BP.

XX

AC AAI28684;

XX

DT 12-OCT-2001 (first entry)

XX

DE Colon tumour related determined cDNA sequence for clone 25275.

XX

KW Human; immunotherapy; diagnosis; colon cancer; colon tumour; immunogenic;  
 KW gene therapy; vaccine; colonic cancer; ss.

XX

OS Homo sapiens.

XX

PN WO200149716-A2.

XX

PD 12-JUL-2001.

XX

PF 29-DEC-2000; 2000WO-US035596.

XX

PR 30-DEC-1999; 99US-00476296.

PR 10-JAN-2000; 2000US-00480321.

PR 15-FEB-2000; 2000US-00504629.

PR 06-MAR-2000; 2000US-00519444.

PR 19-MAY-2000; 2000US-00575251.

PR 29-JUN-2000; 2000US-00609448.

PR 28-AUG-2000; 2000US-00649811.

XX

PA (CORI-) CORIXA CORP.

XX

PI Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;

QV 552 AGTGGCAGTCGCATTGTCTCTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG 611

```

      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 AGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG 300
Qy      612 ATACCCTGCTTTGGGTTTGTAAAGTTTTGCACTGTAGGGTTTTGTGGAATTGGGAGCCT 671
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 ATACCCTGCTTTGGGTTTGTAAAGTTTTGCACTGTAGGGTTTTGTGGAATTGGGAGCCT 360
Qy      672 AATTGATTTTCATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 731
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 AATTGATTTTCATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 420
Qy      732 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 791
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 480
Qy      792 AACGCAATTATATCCATAA 810
      ||||||||||||||||||
Db      481 AACGCAATTATATCCATAA 499

```

RESULT 9

AAI28696

ID AAI28696 standard; cDNA; 508 BP.

XX

AC AAI28696;

XX

DT 12-OCT-2001 (first entry)

XX

DE Colon tumour related determined cDNA sequence for clone 25288.

XX

KW Human; immunotherapy; diagnosis; colon cancer; colon tumour; immunogenic;  
KW gene therapy; vaccine; colonic cancer; ss.

XX

OS Homo sapiens.

XX

PN WO200149716-A2.

XX

PD 12-JUL-2001.

XX

PF 29-DEC-2000; 2000WO-US035596.

XX

PR 30-DEC-1999; 99US-00476296.

PR 10-JAN-2000; 2000US-00480321.

PR 15-FEB-2000; 2000US-00504629.

PR 06-MAR-2000; 2000US-00519444.

PR 19-MAY-2000; 2000US-00575251.

PR 29-JUN-2000; 2000US-00609448.

PR 28-AUG-2000; 2000US-00649811.

XX

PA (CORI-) CORIXA CORP.

XX

PI Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;

PI King GE, Wang T, Jiang Y;

XX

DR WPI; 2001-441847/47.

XX

PT Colon tumor associated proteins and nucleic acids useful for the

PT prevention, diagnosis and treatment of colonic cancer.

XX

PS Claim 2; Page 201; 472pp; English.

XX

CC The present invention describes colon tumour associated proteins (I) and  
CC the polynucleotides (II) that encode them. (I) have cytostatic activity.  
CC (I) and (II) can be used in gene therapy and vaccine production. (I) and  
CC (II) may be used in the prevention, diagnosis and treatment of diseases  
CC associated with inappropriate colon tumour associated protein (TCAP)  
CC expression, such as colonic cancer. For example, (I) and (II) may be used  
CC to treat disorders associated with decreased expression by rectifying  
CC mutations or deletions in a patient's genome that affect the activity of  
CC TCAPs by expressing inactive proteins or to supplement the patients own  
CC production of them. Additionally, (II) may be used to produce the TCAP  
CC proteins, by inserting the nucleic acids into a host cell culturing the  
CC cell to express the protein. (II) and its complementary sequences may  
CC also be used as DNA probes in diagnostic polymerase chain reaction (PCR)  
CC and hybridisation assays to detect and quantitate the presence of similar  
CC nucleic acids in samples, and therefore which patients may be in need of  
CC restorative therapy. (I) may also be used as antigens in the production  
CC of antibodies against TCAPs and in assays to identify modulators of TCAP  
CC expression and activity. Anti-(I) antibodies and antagonists may also be  
CC used to down regulate TCAP expression and activity. The anti-(I)  
CC antibodies may also be used as diagnostic agents for detecting the  
CC presence of TCAPs in samples (e.g. by enzyme linked immunosorbant assay  
CC (ELISA)). AAI28460 to AAI29512 and AAM24494 to AAM24523 represent  
CC nucleotide and amino acid sequences given in the exemplification of the  
CC present invention

XX

SQ Sequence 508 BP; 153 A; 89 C; 103 G; 163 T; 0 U; 0 Other;

Query Match 61.6%; Score 499; DB 4; Length 508;  
Best Local Similarity 100.0%; Pred. No. 9.4e-140;  
Matches 499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	312	CGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA	371
Db	1	CGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA	60
Qy	372	AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCTCATGTTTCCTG	431
Db	61	AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCTCATGTTTCCTG	120
Qy	432	TTTTCCAGCACCCAACATAACTTGTAAAGATTCCAGTGGCAATGAAACACATTTTACTGG	491
Db	121	TTTTCCAGCACCCAACATAACTTGTAAAGATTCCAGTGGCAATGAAACACATTTTACTGG	180
Qy	492	GAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCTACAA	551
Db	181	GAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCTACAA	240
Qy	552	AGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG	611
Db	241	AGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG	300
Qy	612	ATACCCTGCTTTGGGTTTGTAAAGTTTGCCTGTAGGGTTTGTGGAATTGGGAGCCT	671

Db 301 ATACCCTGCTTTGGGTTTGTAAAGTTTTGCACTGTAGGGTTTTGTGGAATTGGGAGCCT 360

Qy 672 AATTGATTTTCATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 731  
 |||

Db 361 AATTGATTTTCATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 420

Qy 732 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 791  
 |||

Db 421 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 480

Qy 792 AACGCAATTATATCCATAA 810  
 |||

Db 481 AACGCAATTATATCCATAA 499

RESULT 10

ABZ32882

ID ABZ32882 standard; cDNA; 508 BP.

XX

AC ABZ32882;

XX

DT 30-JAN-2003 (first entry)

XX

DE Human colon tumour cDNA clone 25288 SEQ ID NO:245.

XX

KW Human; colon cancer; colon tumour; immunotherapy; diagnosis; cancer;

KW tumour; immune response; immunostimulant; cytostatic; vaccine; gene; ss.

XX

OS Homo sapiens.

XX

PN WO200283070-A2.

XX

PD 24-OCT-2002.

XX

PF 09-APR-2002; 2002WO-US011475.

XX

PR 10-APR-2001; 2001US-00833263.

PR 03-AUG-2001; 2001US-00922217.

PR 19-DEC-2001; 2001US-00025380.

XX

PA (CORI-) CORIXA CORP.

XX

PI Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;

PI Wang T, Jiang Y, Smith CL, King GE, Wang A, Clapper JD, Skeiky YAW;

PI Fanger GR, Vedvick TS, Carter D;

XX

DR WPI; 2003=067548/06.

XX

PT New polynucleotide, useful for the preparation of a composition for  
 PT stimulating an immune response against, or treating, cancer.

XX

PS Example 1; Page 204; 537pp; English.

XX

CC The present invention describes compounds (I) for the immunotherapy and  
 CC diagnosis of colon cancer. Also described: (1) a method for detecting the  
 CC presence of cancer in a patient; (2) a method for stimulating and/or  
 CC expanding T cells specific for a tumour protein; (3) an isolated T cell

CC population comprising T cells prepared by the method of (2); (4) a method  
CC for stimulating an immune response in a patient; (5) a method for  
CC treating cancer in a patient; and (6) a method for inhibiting the  
CC development of cancer in a patient. (I) have immunostimulant and  
CC cytostatic activities and can be used in vaccines. ABZ32646 to ABZ33725  
CC and ABP55343 to ABP55391 represent human colon cancer/tumour related  
CC sequences used in the exemplification of the present invention

XX

SQ Sequence 508 BP; 153 A; 89 C; 103 G; 163 T; 0 U; 0 Other;

Query Match 61.6%; Score 499; DB 7; Length 508;  
Best Local Similarity 100.0%; Pred. No. 9.4e-140;  
Matches 499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      312 CGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 371
          |||
Db      1   CGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 60

Qy      372 AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACACAGCTCATGTTTCCTG 431
          |||
Db      61 AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACACAGCTCATGTTTCCTG 120

Qy      432 TTTTCCAGCACCCAACATAAAGTGTAAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 491
          |||
Db      121 TTTTCCAGCACCCAACATAAAGTGTAAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 180

Qy      492 GAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCCTACAA 551
          |||
Db      181 GAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCCTACAA 240

Qy      552 AGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG 611
          |||
Db      241 AGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG 300

Qy      612 ATACCCTGCTTTGGGTTTGTAAAGTTTTGCACTGTAGGGTTTTGTGGAATTGGGAGCCT 671
          |||
Db      301 ATACCCTGCTTTGGGTTTGTAAAGTTTTGCACTGTAGGGTTTTGTGGAATTGGGAGCCT 360

Qy      672 AATTGATTTTCATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 731
          |||
Db      361 AATTGATTTTCATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 420

Qy      732 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 791
          |||
Db      421 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 480

Qy      792 AACGCAATTATATCCATAA-810
          |||
Db      481 AACGCAATTATATCCATAA 499
```

RESULT 11

ABZ32870

ID ABZ32870 standard; cDNA; 508 BP.

XX

AC ABZ32870;

XX



DT 30-JAN-2003 (first entry)  
 XX  
 DE Human colon tumour cDNA clone 25275 SEQ ID NO:233.  
 XX  
 KW Human; colon cancer; colon tumour; immunotherapy; diagnosis; cancer;  
 KW tumour; immune response; immunostimulant; cytostatic; vaccine; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200283070-A2.  
 XX  
 PD 24-OCT-2002.  
 XX  
 PF 09-APR-2002; 2002WO-US011475.  
 XX  
 PR 10-APR-2001; 2001US-00833263.  
 PR 03-AUG-2001; 2001US-00922217.  
 PR 19-DEC-2001; 2001US-00025380.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;  
 PI Wang T, Jiang Y, Smith CL, King GE, Wang A, Clapper JD, Skeiky YAW;  
 PI Fanger GR, Vedvick TS, Carter D;  
 XX  
 DR WPI; 2003-067548/06.  
 XX  
 PT New polynucleotide, useful for the preparation of a composition for  
 PT stimulating an immune response against, or treating, cancer.  
 XX  
 PS Example 1; Page 201; 537pp; English.  
 XX  
 CC The present invention describes compounds (I) for the immunotherapy and  
 CC diagnosis of colon cancer. Also described: (1) a method for detecting the  
 CC presence of cancer in a patient; (2) a method for stimulating and/or  
 CC expanding T cells specific for a tumour protein; (3) an isolated T cell  
 CC population comprising T cells prepared by the method of (2); (4) a method  
 CC for stimulating an immune response in a patient; (5) a method for  
 CC treating cancer in a patient; and (6) a method for inhibiting the  
 CC development of cancer in a patient. (I) have immunostimulant and  
 CC cytostatic activities and can be used in vaccines. ABZ32646 to ABZ33725  
 CC and ABP55343 to ABP55391 represent human colon cancer/tumour related  
 CC sequences used in the exemplification of the present invention  
 XX  
 SQ Sequence 508 BP; 153 A; 89 C; 103 G; 163 T; 0 U; 0 Other;

---

Query Match 61.6%; Score 499; DB 7; Length 508;  
 Best Local Similarity 100.0%; Pred. No. 9.4e-140;  
 Matches 499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	312	CGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA	371
Db	1	CGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA	60
Qy	372	AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACACAGCTCATGTTTCCTG	431
Db	61	AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACACAGCTCATGTTTCCTG	120

Qy 432 TTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 491  
 |||  
 Db 121 TTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 180  
 Qy 492 GAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCCTACAA 551  
 |||  
 Db 181 GAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCCTACAA 240  
 Qy 552 AGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG 611  
 |||  
 Db 241 AGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG 300  
 Qy 612 ATACCCTGCTTTGGGTTTGTTAAAGTTTTGCACTGTAGGGTTTTGTGGAATTGGGAGCCT 671  
 |||  
 Db 301 ATACCCTGCTTTGGGTTTGTTAAAGTTTTGCACTGTAGGGTTTTGTGGAATTGGGAGCCT 360  
 Qy 672 AATTGATTTTCATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 731  
 |||  
 Db 361 AATTGATTTTCATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 420  
 Qy 732 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 791  
 |||  
 Db 421 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 480  
 Qy 792 AACGCAATTATATCCATAA 810  
 |||  
 Db 481 AACGCAATTATATCCATAA 499

# RESULT 12

ABK52558

ID ABK52558 standard; cDNA; 1095 BP.

XX

AC ABK52558;

XX

DT 13-AUG-2002 (first entry)

XX

DE cDNA encoding RNA polymerase II subunit 11.

XX

KW RNA polymerase II subunit 11; ss; gene; cancer; HIV; infection;

KW human immunodeficiency virus.

XX

OS Unidentified.

XX

FH Key Location/Qualifiers

FT CDS 12..314

FT /\*tag= a

FT /product= "RNA polymerase II subunit 11"

XX

PN CN1331300-A.

XX

PD 16-JAN-2002.

XX

PF 30-JUN-2000; 2000CN-00116963.

XX

PR 30-JUN-2000; 2000CN-00116963.



Qy	496	GAAGTTGGTTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCTTACAAAGTG	555
Db	422	GAAGTTGGTTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCTTACAAAGTG	481
Qy	556	GCAGTCGCATTGTCTCTTTTTCTTGGATGGTGGGAGCAGATCGATTTTACCTTGGATAC	615
Db	482	GCAGTCGCATTGTCTCTTTTTCTTGGATGGTGGGAGCAGATCGATTTTACCTTGGATAC	541
Qy	616	CCTGCTTTGGGTTTGTAAAGTTTTGCACTGTAGGGTTTGTGGAATTGGGAGCCTAATT	675
Db	542	CCTGCTTTGGGTTTGTAAAGTTTTGCACTGTAGGGTTTGTGGAATTGGGAGCCTAATT	601
Qy	676	GATTTCAATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACATTATA	735
Db	602	GATTTCAATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACATTATA	661
Qy	736	GATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAAAACG	795
Db	662	GATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAAAACG	721
Qy	796	CAATTATATCCATAA	810
Db	722	CAATTATATCCATAA	736

# RESULT 13

AAX41191

ID AAX41191 standard; cDNA; 440 BP.

XX

AC AAX41191;

XX

DT 17-JUN-1999 (first entry)

XX

DE Human secreted protein 5' EST SEQ ID NO:135.

XX

KW Human; secreted protein; EST; expressed sequence tag; diagnosis;

KW forensic; gene therapy; chromosome mapping; signal peptide;

KW upstream regulatory sequence; cytokine activity; cell proliferation;

KW differentiation; haematopoiesis regulation; tissue growth regulation;

KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;

KW thrombolytic; anti-inflammatory; tumour inhibition; ds.

XX

OS Homo sapiens.

XX

PN WO9906548-A2.

XX

PD 11-FEB-1999.

XX

PF 31-JUL-1998; 98WO-IB001222.

XX

PR 01-AUG-1997; 97US-00905135.

XX

PA (GEST ) GENSET.

XX

PI Dumas Milne Edwards J, Duclert A, Lacroix B;

XX

DR WPI; 1999-153778/13.

DR P-PSDB; AAY12358.  
 XX  
 PT New nucleic acids encoding human secreted proteins - obtained from cDNA  
 PT libraries prepared from e.g. liver, ovary, brain, prostate, kidney, lung,  
 PT umbilical cord, placenta and colon tissue.  
 XX  
 PS Claim 1; Page 315; 824pp; English.  
 XX  
 CC AAX41094 to AAX41347 represent 5' expressed sequence tags (ESTs) for  
 CC human secreted proteins, and encode the proteins given in AAY12261 to  
 CC AAY12514, respectively. The proteins given represent the signal peptide  
 CC and an N-terminal fragment of a secreted protein. The nucleic acid  
 CC sequences can be used for producing secreted human gene products. They  
 CC can also be used to develop products for diagnosis and therapy. The  
 CC proteins obtained may have cytokine activity, cell  
 CC proliferation/differentiation activity, haematopoiesis regulating  
 CC activity, tissue growth regulating activity, reproductive hormone  
 CC regulating activity, chemotactic/ chemokinetic activity, haemostatic and  
 CC thrombolytic activity, receptor/ ligand activity, anti-inflammatory  
 CC activity, tumour inhibition activity or other activities. The products  
 CC can be used in forensic, gene therapy and chromosome mapping procedures.  
 CC The sequences can also be used for obtaining corresponding promoter  
 CC sequences. The nucleic acids encoding the signal peptide can be used for  
 CC directing extracellular secretion of a polypeptide or the insertion of a  
 CC polypeptide into a membrane, or importing a polypeptide into a cell  
 XX  
 SQ Sequence 440 BP; 107 A; 103 C; 114 G; 114 T; 0 U; 2 Other;

Query Match 53.9%; Score 436.8; DB 2; Length 440;  
 Best Local Similarity 99.5%; Pred. No. 4.9e-121;  
 Matches 436; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy	166	GAGAAAGTGTCTGGTCTCCAAGATGGCGGCCGCTGGCCGTCTGGTCCGTCTGCTCCGGAG	225
Db	3	GAGAAAGTGTCTGGTCTCCAAGATGGCGGCCGCTGGCSDTCTGGTCCGTCTGCTCCGGAG	62
Qy	226	GCCGTGACGGCCAGACTCGTTGGTGTCTGTGGTTTCGTCTCAGTCACTACAGGACCCTGG	285
Db	63	GCCGTGACGGCCAGACTCGTTGGTGTCTGTGGTTTCGTCTCAGTCACTACAGGACCCTGG	122
Qy	286	GGGGCTGTTGCCACCTCCGCCGGGGGCGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTG	345
Db	123	GGGGCTGTTGCCACCTCCGCCGGGGGCGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTG	182
Qy	346	GGACAATATATTTGTAAAGATCCAAAAATAAATGACGCTACGCAAGAACCAGTTAACTGT	405
Db	183	GGACAATATATTTGTAAAGATCCAAAAATAAATGACGCTACGCAAGAACCAGTTAACTGT	242
Qy	406	ACAAACTACACAGCTCATGTTTCCTGTTTTCAGCACCCAACATAACTTGTAAGGATTCC	465
Db	243	ACAAACTACACAGCTCATGTTTCCTGTTTTCAGCACCCAACATAACTTGTAAGGATTCC	302
Qy	466	AGTGGCAATGAAACACATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGC	525
Db	303	AGTGGCAATGAAACACATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGC	362
Qy	526	CGAAATGTAAATGGCTATTCCTACAAAGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGG	585

```

          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      363 CGAAATGTAAATGGCTATTCCTACAAAGTGGCAGTCGCATTGTCTCTTTTCTTGGATGG 422
Qy      586 TTGGGAGCAGATCGATTT 603
          ||||||||||||||||
Db      423 TTGGGAGCAGATCGATTT 440

```

RESULT 14

AAX41259

ID AAX41259 standard; cDNA; 455 BP.

XX

AC AAX41259;

XX

DT 17-JUN-1999 (first entry)

XX

DE Human secreted protein 5' EST SEQ ID NO:203.

XX

KW Human; secreted protein; EST; expressed sequence tag; diagnosis;  
KW forensic; gene therapy; chromosome mapping; signal peptide;  
KW upstream regulatory sequence; cytokine activity; cell proliferation;  
KW differentiation; haematopoiesis regulation; tissue growth regulation;  
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;  
KW thrombolytic; anti-inflammatory; tumour inhibition; ds.

XX

OS Homo sapiens.

XX

PN WO9906548-A2.

XX

PD 11-FEB-1999.

XX

PF 31-JUL-1998; 98WO-IB001222.

XX

PR 01-AUG-1997; 97US-00905135.

XX

PA (GEST ) GENSET.

XX

PI Dumas Milne Edwards J, Duclert A, Lacroix B;

XX

DR WPI; 1999-153778/13.

DR

P-PSDB; AAY12426.

XX

PT New nucleic acids encoding human secreted proteins - obtained from cDNA  
PT libraries prepared from e.g. liver, ovary, brain, prostate, kidney, lung,  
PT umbilical cord, placenta and colon tissue.

XX

PS Claim 1; Page 456; 824pp; English.

XX

CC AAX41094 to AAX41347 represent 5' expressed sequence tags (ESTs) for  
CC human secreted proteins, and encode the proteins given in AAY12261 to  
CC AAY12514, respectively. The proteins given represent the signal peptide  
CC and an N-terminal fragment of a secreted protein. The nucleic acid  
CC sequences can be used for producing secreted human gene products. They  
CC can also be used to develop products for diagnosis and therapy. The  
CC proteins obtained may have cytokine activity, cell  
CC proliferation/differentiation activity, haematopoiesis regulating  
CC activity, tissue growth regulating activity, reproductive hormone

CC regulating activity, chemotactic/ chemokinetic activity, haemostatic and  
 CC thrombolytic activity, receptor/ ligand activity, anti-inflammatory  
 CC activity, tumour inhibition activity or other activities. The products  
 CC can be used in forensic, gene therapy and chromosome mapping procedures.  
 CC The sequences can also be used for obtaining corresponding promoter  
 CC sequences. The nucleic acids encoding the signal peptide can be used for  
 CC directing extracellular secretion of a polypeptide or the insertion of a  
 CC polypeptide into a membrane, or importing a polypeptide into a cell

XX

SEQ Sequence 455 BP; 102 A; 107 C; 115 G; 122 T; 0 U; 9 Other;

Query Match 52.5%; Score 425.2; DB 2; Length 455;  
 Best Local Similarity 96.3%; Pred. No. 1.6e-117;  
 Matches 439; Conservative 3; Mismatches 12; Indels 2; Gaps 1;

Qy	177	GGTCTCCAAGATGGCGGCCGCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGGC	236
		:	
Db	2	GGTCTCCAAGATGGCGGCCGCTGGCCGTCTGGTCCGKCTGCTCCGGAGGCCGTGACGGC	61
Qy	237	CAGACTCGTTGGTGTCTGTGGTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGC	296
Db	62	CAGACTCGTTGGTGTCTGTGGTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGC	121
Qy	297	CACCTCCGCCGGGGGCGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATAT	356
Db	122	CACCTCCGCCGGGGGCGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATAT	181
Qy	357	TTGTAAAGATCCAAAAATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACAC	416
Db	182	TTGTAAAGATCCAAAAATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACAC	241
Qy	417	AGCTCATGTTTCTGTGTTTCCAGCACCCCAACATAACTTGTAAGGATTCCAGTGGCAATGA	476
Db	242	AGCTCATGTTTCTGTGTTTCCAGCACCCCAACATAACTTGTAAGGATNCCAGTGGCAATGA	301
Qy	477	AACACATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAA	536
Db	302	AACACATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAA	361
Qy	537	TGGCTATTCCTACAAAGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGA	596
		: :	
Db	362	TGGCTATTCCTAC--NNTKAGCAGTNNNWTGTCTCTTTTTCTTGGATGGTTGGGAGCAGA	419
Qy	597	TCGATTTTACCTTGGATACCCTGCTTTGGGTTTGTT	632
Db	420	TCGATTTTACCTTGGATACCCTGCTTTGGGTTTGTT	455

# RESULT 15

AAC04131

ID AAC04131 standard; cDNA; 487 BP.

XX

AC AAC04131;

XX

DT 06-OCT-2000 (first entry)

XX

DE Human secreted protein 5' EST, SEQ ID NO: 8206.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping; ss.  
XX  
OS Homo sapiens.  
XX  
PN EP1033401-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 21-FEB-2000; 2000EP-00200610.  
XX  
PR 26-FEB-1999; 99US-0122487P.  
XX  
PA (GEST ) GENSET.  
XX  
PI Dumas Milne Edwards J, Duclert A, Giordano J;  
XX  
DR WPI; 2000-500381/45.  
XX  
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.  
XX  
PS Claim 1; SEQ ID NO 8206; 71pp + Sequence Listing; English.  
XX  
CC The present sequence is one of a large number of 5' ESTs derived from  
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively  
CC identified within the present sequence. The 5' ESTs were prepared from  
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST  
CC sequences usually correspond mainly to the 3' untranslated region (UTR)  
CC of the mRNA because they are often obtained from oligo-dT primed cDNA  
CC libraries. Such ESTs are not well suited for isolating cDNA sequences  
CC derived from the 5' ends of mRNAs and even in those cases where longer  
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'  
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used  
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in  
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.  
CC They are used to obtain upstream regulatory sequences and to design  
CC expression and secretion vectors  
XX  
SQ Sequence 487 BP; 115 A; 118 C; 125 G; 120 T; 0 U; 9 Other;

QY	164	GCGAGAAAGTGTCTGGTCTCCAAGATGGCGGCCGCCTGGCCGTCTGGTCCGTCTGCTCCGG	223
Db	1	GCGAGAAAGTGTCTGGTCTCCAAGATGGCGGCCGCM	60
QY	224	AGGCCGTGACGGCCAGACTCGTTGGTGTCTGTGGTTCGTCTCAGTCACTACAGGACCCT	283
Db	61	AAGCCGTGACGGCCAGAMTCGTTGGTGTCTGTGGTTCGTMTCACTACAGGACCCT	120
QY	284	GGGGGGCTGTTGCCACCTCCGCCGGGGGCGAGGAGTCGCTTAAGTGGCAGGACCTCAAAG	343





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2004, 06:47:39 ; Search time 91 Seconds  
(without alignments)  
4939.673 Million cell updates/sec

Title: US-09-852-100B-1  
Perfect score: 810  
Sequence: 1 atgcatatttttaaagggtc.....aaacgcaattatatccataa 810

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	499	61.6	508	4	US-09-401-064-233	Sequence 233, App
2	499	61.6	508	4	US-09-401-064-245	Sequence 245, App
3	49.8	6.1	1455	3	US-09-276-531-33	Sequence 33, Appl
4	40.4	5.0	1119	4	US-09-489-039A-6022	Sequence 6022, Ap
5	38.6	4.8	392000	4	US-10-027-983-11	Sequence 11, Appl
6	36.8	4.5	4403765	3	US-09-103-840A-2	Sequence 2, Appli
7	36.4	4.5	1462	1	US-08-552-142A-16	Sequence 16, Appl
8	36.4	4.5	1951	1	US-08-910-973-16	Sequence 16, Appl
9	36.4	4.5	1951	4	US-09-499-227-16	Sequence 16, Appl
10	36.2	4.5	8093	4	US-10-204-708-32	Sequence 32, Appl
11	35.4	4.4	450	4	US-09-252-991A-12127	Sequence 12127, A

c	12	35.4	4.4	1404	4	US-09-252-991A-12291	Sequence 12291, A
	13	35.4	4.4	9347	4	US-10-204-708-36	Sequence 36, Appl
	14	35.4	4.4	580073	4	US-08-545-528D-1	Sequence 1, Appli
	15	35.2	4.3	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c	16	35	4.3	1494	4	US-09-252-991A-7049	Sequence 7049, Ap
	17	35	4.3	4236	4	US-09-252-991A-7057	Sequence 7057, Ap
	18	35	4.3	7304	4	US-10-204-708-43	Sequence 43, Appl
c	19	35	4.3	10023	4	US-09-252-991A-6997	Sequence 6997, Ap
	20	35	4.3	1830121	4	US-09-557-884-1	Sequence 1, Appli
	21	35	4.3	1830121	4	US-09-643-990A-1	Sequence 1, Appli
	22	34.8	4.3	832	4	US-09-621-976-2813	Sequence 2813, Ap
c	23	34.6	4.3	4673	1	US-07-638-431-1	Sequence 1, Appli
c	24	34.6	4.3	4673	5	PCT-US92-00018-1	Sequence 1, Appli
	25	34	4.2	5152	4	US-10-204-708-47	Sequence 47, Appl
	26	34	4.2	11131	4	US-10-204-708-27	Sequence 27, Appl
	27	33.6	4.1	11049	4	US-10-204-708-23	Sequence 23, Appl
c	28	33.4	4.1	549	4	US-09-252-991A-14907	Sequence 14907, A
	29	33.4	4.1	1125	4	US-09-252-991A-14723	Sequence 14723, A
	30	33.4	4.1	1636	6	5447867-2	Patent No. 5447867
c	31	33.4	4.1	1854	4	US-09-252-991A-15029	Sequence 15029, A
	32	33.4	4.1	9179	4	US-08-956-171E-100	Sequence 100, App
c	33	33.4	4.1	319608	4	US-09-539-333D-1	Sequence 1, Appli
c	34	33.4	4.1	319608	4	US-09-679-409-1	Sequence 1, Appli
c	35	33	4.1	364	4	US-09-621-976-17202	Sequence 17202, A
c	36	33	4.1	3255	4	US-09-601-198-108	Sequence 108, App
	37	33	4.1	4029	4	US-09-620-312D-201	Sequence 201, App
	38	33	4.1	5152	4	US-10-204-708-48	Sequence 48, Appl
c	39	33	4.1	55298	4	US-09-491-356C-1	Sequence 1, Appli
c	40	32.8	4.0	988	1	US-08-243-545-5	Sequence 5, Appli
c	41	32.8	4.0	988	2	US-08-993-962-5	Sequence 5, Appli
c	42	32.8	4.0	988	3	US-09-160-841-5	Sequence 5, Appli
c	43	32.8	4.0	988	3	US-09-109-100-2	Sequence 2, Appli
c	44	32.8	4.0	988	4	US-08-669-692-5	Sequence 5, Appli
c	45	32.8	4.0	988	4	US-08-444-626-5	Sequence 5, Appli

#### ALIGNMENTS

#### RESULT 1

US-09-401-064-233

; Sequence 233, Application US/09401064

; Patent No. 6623923

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Lodes, Michael J.

; APPLICANT: Secrist, Heather

; APPLICANT: Benson, Darin R.

; APPLICANT: Meagher, Madeline Joy

; APPLICANT: Stolk, John A.

; APPLICANT: Wang, Tongtong

; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND

; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE

; FILE REFERENCE: 210121.471C2

; CURRENT APPLICATION NUMBER: US/09/401,064

; CURRENT FILING DATE: 1999-09-22

; NUMBER OF SEQ ID NOS: 371

; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 233  
; LENGTH: 508  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-401-064-233

Query Match 61.6%; Score 499; DB 4; Length 508;  
Best Local Similarity 100.0%; Pred. No. 2.1e-153;  
Matches 499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      312 CGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 371
          |||
Db       1 CGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 60

Qy      372 AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACACAGCTCATGTTTCCTG 431
          |||
Db      61 AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACACAGCTCATGTTTCCTG 120

Qy      432 TTTTCCAGCACCCAACATAAAGTGTAAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 491
          |||
Db     121 TTTTCCAGCACCCAACATAAAGTGTAAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 180

Qy      492 GAACGAAGTTGGTTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCTCTACAA 551
          |||
Db     181 GAACGAAGTTGGTTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCTCTACAA 240

Qy      552 AGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG 611
          |||
Db     241 AGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG 300

Qy      612 ATACCCTGCTTTGGGTTTGTAAAGTTTTGCACTGTAGGGTTTTGTGGAATTGGGAGCCT 671
          |||
Db     301 ATACCCTGCTTTGGGTTTGTAAAGTTTTGCACTGTAGGGTTTTGTGGAATTGGGAGCCT 360

Qy      672 AATTGATTTTCAATCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 731
          |||
Db     361 AATTGATTTTCAATCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 420

Qy      732 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 791
          |||
Db     421 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 480

Qy      792 AACGCAATTATATCCATAA 810
          |||
Db     481 AACGCAATTATATCCATAA 499
```

RESULT 2

US-09-401-064-245

; Sequence 245, Application US/09401064  
; Patent No. 6623923  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Benson, Darin R.

```
; APPLICANT: Meagher, Madeline Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C2
; CURRENT APPLICATION NUMBER: US/09/401,064
; CURRENT FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 371
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 245
; LENGTH: 508
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-401-064-245
```

```
Query Match          61.6%; Score 499; DB 4; Length 508;
Best Local Similarity 100.0%; Pred. No. 2.1e-153;
Matches 499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      312 CGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 371
          |||
Db       1 CGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 60

Qy      372 AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACACAGCTCATGTTTCCTG 431
          |||
Db       61 AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACACAGCTCATGTTTCCTG 120

Qy      432 TTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 491
          |||
Db      121 TTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 180

Qy      492 GAACGAAGTTGGTTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTTCCTACAA 551
          |||
Db      181 GAACGAAGTTGGTTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTTCCTACAA 240

Qy      552 AGTGGCAGTCGCATTGTCTCTTTTCTTGATGGTTGGGAGCAGATCGATTTTACCTTGG 611
          |||
Db      241 AGTGGCAGTCGCATTGTCTCTTTTCTTGATGGTTGGGAGCAGATCGATTTTACCTTGG 300

Qy      612 ATACCCTGCTTTGGGTTTGTAAAGTTTTGCACTGTAGGGTTTTGTGGAATTGGGAGCCT 671
          |||
Db      301 ATACCCTGCTTTGGGTTTGTAAAGTTTTGCACTGTAGGGTTTTGTGGAATTGGGAGCCT 360

Qy      672 AATTGATTTTATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 731
          |||
Db      361 AATTGATTTTATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 420

Qy      732 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 791
          |||
Db      421 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 480

Qy      792 AACGCAATTATATCCATAA 810
          |||
Db      481 AACGCAATTATATCCATAA 499
```

RESULT 3

US-09-276-531-33

; Sequence 33, Application US/09276531

; Patent No. 6183968

; GENERAL INFORMATION:

; APPLICANT: Bandman, Olga

; APPLICANT: Lal, Preeti

; APPLICANT: Hillman, Jennifer L.

; APPLICANT: Yue, Henry

; APPLICANT: Reddy, Roopa

; APPLICANT: Guegler, Karl J.

; APPLICANT: Baughn, Mariah R.

; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING

; TITLE OF INVENTION: RECEPTORS AND PROTEINS ASSOCIATED WITH CELL

PROLIFERATION

; NUMBER OF SEQUENCES: 134

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

; STREET: 3174 PORTER DRIVE

; CITY: PALO ALTO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/276,531

; FILING DATE: Herewith

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/079,677

; FILING DATE: March 27, 1998

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Lynn E. Murry, Ph.D.

; REGISTRATION NUMBER: 42,918

; REFERENCE/DOCKET NUMBER: PA-0008 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (650) 855-0555

; TELEFAX: (650) 845-4166

; INFORMATION FOR SEQ ID NO: 33:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1455 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: BRAITUT01

; CLONE: 746308

US-09-276-531-33

Query Match 6.1%; Score 49.8; DB 3; Length 1455;

Best Local Similarity 51.1%; Pred. No. 7.8e-06;

Matches 117; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

Qy	105	GAACCTCGCCCTGTTGCCCTTCTCCCTCCCGCTCTCTGGGCGGAGGCGGAAGCGGAAGTGG	164
Db	672	GAACAGCGGCGGTTTCCTCGGCGCGCCGACGCCCTGAACAACGGCGTGTCTGGAGAGTAG	731
Qy	165	CGAGAAAGTGTCGGTCTCCAAGATGGCGGCCGCTGGCCGTCTGGTCCGTCTGCTCCGGA	224
Db	732	CGAACCAGCGGCAGCCGCCGCGACGGTCTCTGCCGCCGGCGCCACGCCAACAGCGCCAGT	791
Qy	225	GGCCGTGACGGCCAGACTCGTTGGTGTCTGTGGTTTCGTCTCAGTCACTACAGGACCTTG	284
Db	792	GACCGCGCCTGGCTCCATTCAAGGTAATGTGGCGCCCGCTGCGGCCACCGCCGAGCCGC	851
Qy	285	GGGGGCTGTTGCCACCTCCGCC	306

## RESULT 5

US-10-027-983-11

; Sequence 11, Application US/10027983

; Patent No. 6617162

; GENERAL INFORMATION:

; APPLICANT: Kenneth W. Dobie

; APPLICANT: Mark P. Roach

; TITLE OF INVENTION: ANTISENSE MODULATION OF ESTROGEN RECEPTOR ALPHA  
EXPRESSION

; FILE REFERENCE: RTS-0340

; CURRENT APPLICATION NUMBER: US/10/027,983

; CURRENT FILING DATE: 2001-12-18

; NUMBER OF SEQ ID NOS: 98

; SEQ ID NO 11

; LENGTH: 392000

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: unsure

; LOCATION: 137740

; OTHER INFORMATION: unknown

; NAME/KEY: unsure

; LOCATION: 137742

; OTHER INFORMATION: unknown

; NAME/KEY: misc\_feature

; LOCATION: (138122)...(138221)

; OTHER INFORMATION: n = A,T,C or G

; NAME/KEY: unsure

; LOCATION: 145507

; OTHER INFORMATION: unknown

; NAME/KEY: unsure

; LOCATION: 151967

; OTHER INFORMATION: unknown

; NAME/KEY: misc\_feature

; LOCATION: (151967)...(1542066)

; OTHER INFORMATION: n = A,T,C or G

; NAME/KEY: unsure

; LOCATION: 154217

; OTHER INFORMATION: unknown

; NAME/KEY: misc\_feature

; LOCATION: (164037)...(164136)

; OTHER INFORMATION: n = A,T,C or G

; NAME/KEY: misc\_feature

; LOCATION: (174657)...(174756)

; OTHER INFORMATION: n = A,T,C or G

; NAME/KEY: misc\_feature

; LOCATION: (186224)...(186323)

; OTHER INFORMATION: n = A,T,C or G

; NAME/KEY: misc\_feature

; LOCATION: (195242)...(195341)

; OTHER INFORMATION: n = A,T,C or G

; NAME/KEY: unsure

; LOCATION: 202703

; OTHER INFORMATION: unknown



```

; NAME/KEY: misc_feature
; LOCATION: (202771)...(202870)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc_feature
; LOCATION: (206246)...(215602)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc_feature
; LOCATION: (218126)...(218225)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc_feature
; LOCATION: (220360)...(220459)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc_feature
; LOCATION: (222717)...(222816)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc_feature
; LOCATION: (223981)...(224080)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc_feature
; LOCATION: (227487)...(227586)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc_feature
; LOCATION: (230157)...(230256)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc_feature
; LOCATION: (232299)...(232398)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc_feature
; LOCATION: (236552)...(2366651)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc_feature
; LOCATION: (238789)...(248788)
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: exon
; LOCATION: (118288)...(119101)
; OTHER INFORMATION: exon 1C
; NAME/KEY: exon:intron junction
; LOCATION: (151129)...(151130)
; OTHER INFORMATION: exon 5:intron 5
; NAME/KEY: exon:intron junction
; LOCATION: (299248)...(299249)
; OTHER INFORMATION: exon 9:intron 9
; NAME/KEY: exon:intron junction
; LOCATION: (348578)...(348579)
; OTHER INFORMATION: exon 10:intron 10
; NAME/KEY: intron
; LOCATION: (348579)...(381838)
; OTHER INFORMATION: intron 10
; NAME/KEY: intron:exon junction
; LOCATION: (386185)...(386186)
; OTHER INFORMATION: intron 11:exon 12
US-10-027-983-11

```

```

Query Match          4.8%;  Score 38.6;  DB 4;  Length 392000;
Best Local Similarity 54.6%;  Pred. No. 1.5;
Matches 77;  Conservative 0;  Mismatches 64;  Indels 0;  Gaps 0;

```



Db 676281 ACCCCACGTCCCGGTAGCGCCGTTGCCGCCGTCACCGGGAGCTCCGCCGTCACCGCCGCT  
676340

Qy 292 GTTGCCACCTCCGCCGGGGGCGAGGAGTCGCTTAAGTGCGAGGA 335

|||| | ||||| | || | | || ||| |  
Db 676341 ACCGCCAGCCCCGCCGGCGCCGTGGCTGCCGCCGAGGCCGAGCA 676384

RESULT 7

US-08-552-142A-16

; Sequence 16, Application US/08552142A

; Patent No. 5695995

; GENERAL INFORMATION:

; APPLICANT: Weintraub, Harold M.

; APPLICANT: Lee, Jacqueline E.

; APPLICANT: Tapscott, Stephen J.

; APPLICANT: Hollenberg, Stanley M.

; TITLE OF INVENTION: Neurogenic Differentiation (NeuroD) Genes

; TITLE OF INVENTION: and Proteins

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Christensen O'Connor Johnson Kindness PLLC

; STREET: 1420 Fifth Avenue, Suite 2800

; CITY: Seattle

; STATE: WA

; COUNTRY: USA

; ZIP: 98101-2347

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/552,142A

; FILING DATE: 02-NOV-1995

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/239,238

; FILING DATE: 06-MAY-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: WO PCT/US95/05741

; FILING DATE: 08-MAY-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Broderick, Thomas F.

; REGISTRATION NUMBER: 31,332

; REFERENCE/DOCKET NUMBER: FHCR-1-8933

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 206-682-8100

; TELEFAX: 206-225-0709

; INFORMATION FOR SEQ ID NO: 16:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1462 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; ORIGINAL SOURCE:

; ORGANISM: Mus musculus  
; IMMEDIATE SOURCE:  
; CLONE: 1.1.1  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 231..1101  
US-08-552-142A-16

Query Match 4.5%; Score 36.4; DB 1; Length 1462;  
Best Local Similarity 49.0%; Pred. No. 0.19;  
Matches 97; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

```
Qy      121 CCCTTCTCCCTCCCCTCCTGGGCGGAGGCGGAAGCGGAAGTGGCGAGAAAGTGTCTGGGTC 180
          | | ||| | | | | | | | | | | | | | | | | | |
Db      261 CTCCTCTCGGACGTGCCCAAGTTCGCCAGCTGGGGCGACGGCGACGACGACGAGCCGAGG 320

Qy      181 TCCAAGATGGCGGCCCGCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGGCCAGA 240
          | | | | | | | | | | | | | | | | | | | | | |
Db      321 AGCGACAAGGGCGACGCGCCGCCGAGCCTTCTCCTGCTCCCGGGTCTGGGGGCTCCAGGA 380

Qy      241 CTCGTTGGTGTCTGTGGTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGCCACC 300
          | | | | | | | | | | | | | | | | | | | | | |
Db      381 CCCGCCGGGCGCCAAGCCAGTGTCTCTTCGTGGAGGAGAAGAGATCCCTGAACCCACG 440

Qy      301 TCCGCCGGGGCGAGGAG 318
          | | | | | | | | | |
Db      441 TTGGCTGAGGTCAAGGAG 458
```

RESULT 8

US-08-910-973-16

; Sequence 16, Application US/08910973

; Patent No. 5795723

; GENERAL INFORMATION:

; APPLICANT: Tapscott, Stephen J.

; APPLICANT: Olson, James M.

; TITLE OF INVENTION: Expression of Neurogenic bHLH Genes in Primitive  
Neuroectoder

; NUMBER OF SEQUENCES: 24

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Christensen O'Connor Johnson Kindness PLLC

; STREET: 1420 Fifth Avenue, Suite 2800

; CITY: Seattle

; STATE: WA

; COUNTRY: USA

; ZIP: 98101-2347

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/910,973

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/239,238

```

; FILING DATE: 06-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US95/05741
; FILING DATE: 08-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/17532
; FILING DATE: 30-October-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sheiness, Diana K.
; REGISTRATION NUMBER: 35,356
; REFERENCE/DOCKET NUMBER: FHCR-1-10958
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-682-8100; 206-224-0735 (direct)
; TELEFAX: 206-225-0779
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1951 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Mus musculus
; IMMEDIATE SOURCE:
; CLONE: 1.1.1 (mouse neuroD2)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 230..1378
US-08-910-973-16

```

```

Query Match          4.58; Score 36.4; DB 1; Length 1951;
Best Local Similarity 49.0%; Pred. No. 0.23;
Matches 97; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

```

```

Qy      121 CCCTTCTCCCTCCCGCTCCTGGGCGGAGGCGGAAGCGGAAGTGGCGAGAAAGTGTGCGGTC 180
          | | | | | | | | | | | | | | | | | | | | | | | |
Db      260 CTCCTCTCGGACGTGCCCAAGTTCGCCAGCTGGGGCGACGGCGACGACGAGCCGAGG 319

Qy      181 TCCAAGATGGCGGCCCGCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGGCCAGA 240
          | | | | | | | | | | | | | | | | | | | | | | | |
Db      320 AGCGACAAGGGCGACGCGCCGCCGAGCCTTCTCCTGCTCCCGGGTCGGGGGCTCCAGGA 379

Qy      241 CTCGTTGGTGTCTGTGTTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGCCACC 300
          | | | | | | | | | | | | | | | | | | | | | | | |
Db      380 CCCGCCCCGGGCGCCCAAGCCAGTGTCTCTTCGTGGAGGAGAAGAGATCCCTGAACCCACG 439

-----
Qy      301 TCCGCCGGGGGCGAGGAG-318
          | | | | | | | | | |
Db      440 TTGGCTGAGGTCAAGGAG 457

```

```

RESULT 9
US-09-499-227-16
; Sequence 16, Application US/09499227
; Patent No. 6444463
; GENERAL INFORMATION:
; APPLICANT: Tapscott, Stephen J.

```

```

;   APPLICANT:  Olson, James M.
;   TITLE OF INVENTION:  Expression of Neurogenic bHLH Genes in Primitive
Neuroectoder
;   NUMBER OF SEQUENCES:  24
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE:  Christensen O'Connor Johnson Kindness PLLC
;   STREET:  1420 Fifth Avenue, Suite 2800
;   CITY:  Seattle
;   STATE:  WA
;   COUNTRY:  USA
;   ZIP:  98101-2347
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE:  Floppy disk
;   COMPUTER:  IBM PC compatible
;   OPERATING SYSTEM:  PC-DOS/MS-DOS
;   SOFTWARE:  PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER:  US/09/499,227
;   FILING DATE:  05-August-1998
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:  US 08/239,238
;   FILING DATE:  06-May-1994
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:  WO PCT/US95/05741
;   FILING DATE:  08-May-1995
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:  PCT/US96/17532
;   FILING DATE:  30-October-1996
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:  US 08/910,973
;   FILING DATE:  07-August-1997
;   ATTORNEY/AGENT INFORMATION:
;   NAME:  Sheiness, Diana K.
;   REGISTRATION NUMBER:  35,356
;   REFERENCE/DOCKET NUMBER:  FHCR-1-12742
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE:  206-682-8100; 206-224-0735 (direct)
;   TELEFAX:  206-225-0779
;   INFORMATION FOR SEQ ID NO:  16:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH:  1951 base pairs
;   TYPE:  nucleic acid
;   STRANDEDNESS:  single
;   TOPOLOGY:  linear
;   MOLECULE TYPE:  cDNA
;   ORIGINAL SOURCE:
;   ORGANISM:  Mus musculus
;   IMMEDIATE SOURCE:
;   CLONE:  1.1.1 (mouse neuroD2)
;   FEATURE:
;   NAME/KEY:  CDS
;   LOCATION:  230..1378
US-09-499-227-16

```

```

Query Match          4.5%;  Score 36.4;  DB 4;  Length 1951;
Best Local Similarity 49.0%;  Pred. No. 0.23;
Matches 97;  Conservative 0;  Mismatches 101;  Indels 0;  Gaps 0;

```

```

Qy      121 CCCTTCTCCCTCCCGCTCCTGGGCGGAGGCGGAAGCGGAAGTGGCGAGAAAGTGTCCGGTC 180
      | | | | | | | | | | | | | | | | | | | | | | | |
Db      260 CTCCTCTCGGACGTGCCCAAGTTCGCCAGCTGGGGCGACGGCGACGACGACGAGCCGAGG 319

Qy      181 TCCAAGATGGCGGCCCGCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGGCCAGA 240
      | | | | | | | | | | | | | | | | | | | | | | | |
Db      320 AGCGACAAGGGCGACGCGCCGCCGACGCTTCTCCTGCTCCCGGGTCTGGGGGCTCCAGGA 379

Qy      241 CTCGTTGGTGTCTGTGGTTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGCCACC 300
      | | | | | | | | | | | | | | | | | | | | | | | |
Db      380 CCCGCCCCGGGCCGCCAAGCCAGTGTCTCTTCGTGGAGGAGAAGAGATCCCTGAACCCACG 439

Qy      301 TCCGCCGGGGGCGAGGAG 318
      | | | | | | | | | |
Db      440 TTGGCTGAGGTCAAGGAG 457

```

RESULT 10

US-10-204-708-32

; Sequence 32, Application US/10204708

; Patent No. 6677731

; GENERAL INFORMATION:

; APPLICANT: OLEK, Alexander

; APPLICANT: PIEPENBROCK, Christian

; APPLICANT: BERLIN, Kurt

; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication

; TITLE OF INVENTION: by Assessing DNA Methylation

; FILE REFERENCE: 5013.1012

; CURRENT APPLICATION NUMBER: US/10/204,708

; CURRENT FILING DATE: 2003-05-06

; PRIOR APPLICATION NUMBER: PCT/EP01/03971

; PRIOR FILING DATE: 2001-04-06

; PRIOR APPLICATION NUMBER: DE 10019058.8

; PRIOR FILING DATE: 2000-04-06

; PRIOR APPLICATION NUMBER: DE 10019173.8

; PRIOR FILING DATE: 2000-04-07

; PRIOR APPLICATION NUMBER: DE 10032529.7

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: DE 10043826.1

; PRIOR FILING DATE: 2000-09-01

; NUMBER OF SEQ ID NOS: 98

; SEQ ID NO 32

; LENGTH: 8093

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-204-708-32

Query Match 4.5%; Score 36.2; DB 4; Length 8093;

Best Local Similarity 54.0%; Pred. No. 0.69;

Matches 74; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

```

Qy      673 ATTGATTTTCATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACATT 732
      | | | | | | | | | | | | | | | | | | | | | | | |
Db      6468 AGTAATTTTGTATTTTATTAAAGATAATTTGTTTTGTGTAAAATAGTAATTTTAAATT 6527

```

```

Qy      733 ATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAAA 792
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      6528 TTTGTTTATTATGAAAAGGTAATTTTAAAGTTTATTATGTAAAATTAATTATAAATAGGA 6587

Qy      793 ACGCAATTATATCCATA 809
      | | | | | | | | | |
Db      6588 TTTAATTTATATTTATA 6604

```

RESULT 11

```

US-09-252-991A-12127
; Sequence 12127, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12127
; LENGTH: 450
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12127

```

```

Query Match          4.4%; Score 35.4; DB 4; Length 450;
Best Local Similarity 48.3%; Pred. No. 0.19;
Matches 99; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

```

```

Qy      60 GCGAAGAGACGGAACCTGGCCTCTATCCTATGCGAGGTCCCTTTAAGAACCTCGCCCTGTT 119
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      45 GCGCAGCCAGGTCACCACCTGACCCGCCGCTGCGTCGCGAGGCGCAGGCCGATCCGGT 104

Qy     120 GCCCTTCTCCCTCCCGCTCCTGGGCGGAGGCGGAAGCGGAAGTGGCGAGAAAGTGTCGGT 179
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     105 GCAATTCTCGCAACTGGTCGTGCTTGGCGCGATCGACCGCCTTGGCGGCGACGTCACACC 164

Qy     180 CTCCAAGATGGCGGCCCGCTGGCCGTCTGCTCCGTCTGCTCCGGAGGCCGTGACGGCCAG 239
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     165 TTCCGAGCTGGCCGCCGCCGAGCGGATGCGCTCGTCGAATCTGGCCGCGCTGCTGCGCGA 224

Qy     240 ACTCGTTGGTGTCTGTGGTTCGTC 264
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     225 ACTGGAACGCGGAGGGCTGATCGTC 249

```

RESULT 12

```

US-09-252-991A-12291/c
; Sequence 12291, Application US/09252991A

```



```
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12291
; LENGTH: 1404
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12291
```

```
Query Match          4.4%; Score 35.4; DB 4; Length 1404;
Best Local Similarity 48.3%; Pred. No. 0.4;
Matches 99; Conservative 0; Mismatches 106; Indels 0; Gaps 0;
```

```
Qy      60 GCGAAGAGACGGAACCTGGCCTCTATCCTATGCGAGGTCCCTTTAAGAACCTCGCCCTGTT 119
      ||| || | | || || | | || | | | | | | |
Db      1324 GCGCAGCCAGGTCACCAACCTGACCCGCCGCTGCGTCGCGAGGCGCAGGCCGATCCGGT 1265

Qy      120 GCCCTTCTCCCTCCCGCTCCTGGGCGGAGGCGGAAGCGGAAGTGGCGAGAAAGTGTCCGGT 179
      || ||||| | | | || || | | | | |||| | || |
Db      1264 GCAATTCTCGCAACTGGTCGTGCTTGGCGCGATCGACCGCCTTGGCGGCGACGTCACACC 1205

Qy      180 CTCCAAGATGGCGGCCCGCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGGCCAG 239
      ||| || |||| ||||| || | | |||| | || | || | |
Db      1204 TTCCGAGCTGGCCGCCCGAGCGGATGCGCTCGTCGAATCTGGCCGCGCTGCTGCGCGA 1145

Qy      240 ACTCGTTGGTGTCTGTGGTTTCGTC 264
      ||| | | | | | |||||
Db      1144 ACTGGAACGCGGAGGGCTGATCGTC 1120
```

# RESULT 13

US-10-204-708-36

; Sequence 36, Application US/10204708

; Patent No. 6677731

; GENERAL INFORMATION:

; APPLICANT: OLEK, Alexander

; APPLICANT: PIEPENBROCK, Christian

; APPLICANT: BERLIN, Kurt

; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication

; TITLE OF INVENTION: by Assessing DNA Methylation

; FILE REFERENCE: 5013.1012

; CURRENT APPLICATION NUMBER: US/10/204,708

; CURRENT FILING DATE: 2003-05-06

; PRIOR APPLICATION NUMBER: PCT/EP01/03971

; PRIOR FILING DATE: 2001-04-06

; PRIOR APPLICATION NUMBER: DE 10019058.8

```
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 36
; LENGTH: 9347
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-36
```

```
Query Match          4.4%; Score 35.4; DB 4; Length 9347;
Best Local Similarity 45.7%; Pred. No. 1.4;
Matches 123; Conservative 0; Mismatches 146; Indels 0; Gaps 0;
```

```
Qy      535 AATGGCTATTCCCTACAAAGTGGCAGTCGCATTGTCTCTTTTCTTGGATGGTTGGGAGCA 594
      ||| |  ||  | ||| | | |||  || | || | |  |  ||  ||  |
Db      7248 AATAGTAGTTGAGATAAAATAGAAGTTTATTTTTTTTCGTTAGATAAAATAATTTTAGAGA 7307

Qy      595 GATCGATTTTACCTTGGATACCCTGCTTTGGGTTTGTAAAGTTTGGCACTGTAGGGTTT 654
      |  | |||  |  |  | |||  ||  | | |||  |  |  |||
Db      7308 TAGGTAGTTTAGGGTTAAATGGTGGTTTTACGTTTATTAGGTTTTATCGTTTATTTT 7367

Qy      655 TGTGGAATTGGGAGCCTAATTGATTTTCATTCTTATTTCAATGCAGATTGTTGGACCTTCA 714
      ||  | |  |  ||  ||| ||  ||| ||  ||| ||  ||  |  |  |  |
Db      7368 ATTGTTTTTCGAATTAATTTTATTTTATAGTTATTTTATTGTTTAAGATGGTTTTTGGA 7427

Qy      715 GATGGAAGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACT 774
      |  | || |  ||| ||  ||| ||  |  |  ||  || |  |  |
Db      7428 ATTTTAGTTATTAGATTTATATTTTATTTTAGGTAGCGAAAGTAGGAAGAAGAAAGGGT 7487

Qy      775 AATGAAACATTTAGAAAAACGCAATTATA 803
      ||  |||  ||| ||| |  | || |
Db      7488 AAATAAAATTTTAAAAATATAAAGTAAGA 7516
```

#### RESULT 14

US-08-545-528D-1

; Sequence 1, Application US/08545528D

; Patent No. 6537773

; GENERAL INFORMATION:

; APPLICANT: Fraser et al.

; TITLE OF INVENTION: Nucleotide Sequence of the Mycoplasma Genitalium Genome, Fragments

; Patent No. 6537773

; TITLE OF INVENTION: Thereof, and Uses Thereof

; FILE REFERENCE: PB193P1

; CURRENT APPLICATION NUMBER: US/08/545,528D

; CURRENT FILING DATE: 1995-10-19

; PRIOR APPLICATION NUMBER: US 08/488,018

; PRIOR FILING DATE: 1995-06-07

; PRIOR APPLICATION NUMBER: US 08/473,545

; PRIOR FILING DATE: 1995-06-07  
; NUMBER OF SEQ ID NOS: 1  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 580073  
; TYPE: DNA  
; ORGANISM: Mycoplasma genitalium  
US-08-545-528D-1

Query Match 4.4%; Score 35.4; DB 4; Length 580073;  
Best Local Similarity 55.2%; Pred. No. 21;  
Matches 69; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

Qy 685 CTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACATTATAGATTACTAT 744  
||| ||| || | |||| | || | ||| ||| |

Db 387060 CTTTTTTAAAGTTAAATTGTCTGCTTTGTGAATCCATGGTAAGCAGGATGGAATTATTA  
387119

Qy 745 GGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAAAACGCAATTATAT 804  
| ||| | | |||| || | || | | |||| | || ||| | ||| |

Db 387120 AGTACCTTCCATCCAAGGCTTATTATGAATCATTAACAATAAGCAAATTCAATTTAAAG  
387179

Qy 805 CCATA 809  
|||||  
Db 387180 CCATA 387184

#### RESULT 15

US-09-103-840A-1

; Sequence 1, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; TITLE OF INVENTION: TUBERCULOSIS  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 4411529  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; OTHER INFORMATION: H37Rv  
US-09-103-840A-1

Query Match 4.3%; Score 35.2; DB 3; Length 4411529;  
Best Local Similarity 47.3%; Pred. No. 66;  
Matches 106; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

Qy 112 GCCCTGTTGCCCTTCTCCCTCCCGCTCCTGGGCGGAGGCGGAAGCGGAAGTGGCGAGAAA 171  
||| | || | | || | || | || | ||| | ||| |

Db 674718 GCCGGGCGCGCCGTTTCGCGCCATGCGCGCTGCCGCCGACGCTGGCGCCACCGGCGCCACC  
674777

Qy 172 GTGTCGGTCTCCAAGATGGCGGCCGCCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTG 231

| | | | | | | | | | | | | | | | | | | | | |  
Db 674778 GGCCCCACCGGCGCCCGGGTTGCCGCCATTGCCACCGGTCCCGCCGGCACCAAGGTTGTG  
674837

Qy 232 ACGGCCAGACTCGTTGGTGTCTGTGGTTTCGTCTCAGTCACTACAGGACCCTGGGGGGCT 291

| | | | | | | | | | | | | | | | | | | | | |  
Db 674838 ACCCCACGTCCCGGTAGCGCCGTTGCCGCCGTACCGGGAGCTCCGCCGTACCGCCGCT  
674897

Qy 292 GTTGCCACCTCCGCCGGGGGCGAGGAGTCGCTTAAGTGCGAGGA 335

| | | | | | | | | | | | | | | | | | | | | |  
Db 674898 ACCGCCAGCCCCGCCGGCGCCGTGGCTGCCGCCGAGGCCGAGCA 674941

Search completed: March 4, 2004, 09:18:33  
Job time : 106 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2004, 08:34:04 ; Search time 349 Seconds  
(without alignments)  
8488.596 Million cell updates/sec

Title: US-09-852-100B-1  
Perfect score: 810  
Sequence: 1 atgcatatttttaaagggtc.....aaacgcaattatatccataa 810

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2421054 seqs, 1828716029 residues

Total number of hits satisfying chosen parameters: 4842108

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	IDB	Description
------------	-------	--------------------	-----	-------------

1	810	100.0	810	9	US-09-852-100A-1	Sequence 1, Appli
2	810	100.0	810	9	US-09-833-503A-1	Sequence 1, Appli
3	810	100.0	810	14	US-10-199-881-1	Sequence 1, Appli
4	499	61.6	508	9	US-09-922-217-233	Sequence 233, App
5	499	61.6	508	9	US-09-922-217-245	Sequence 245, App
6	499	61.6	508	9	US-09-833-263-233	Sequence 233, App
7	499	61.6	508	9	US-09-833-263-245	Sequence 245, App
8	499	61.6	508	13	US-10-025-380-233	Sequence 233, App
9	499	61.6	508	13	US-10-025-380-245	Sequence 245, App
10	337.4	41.7	431	10	US-09-918-995-6918	Sequence 6918, Ap
11	100	12.3	256	12	US-10-085-783A-36056	Sequence 36056, A
12	100	12.3	256	15	US-10-242-535A-36056	Sequence 36056, A
13	85.4	10.5	411	12	US-10-085-783A-48351	Sequence 48351, A
14	85.4	10.5	411	15	US-10-242-535A-48351	Sequence 48351, A
15	79	9.8	129	12	US-10-085-783A-16414	Sequence 16414, A
16	79	9.8	129	15	US-10-242-535A-16414	Sequence 16414, A
17	64.6	8.0	161	12	US-10-085-783A-14323	Sequence 14323, A
18	64.6	8.0	161	15	US-10-242-535A-14323	Sequence 14323, A
19	51	6.3	1369	10	US-09-974-879-102	Sequence 102, App
20	51	6.3	1369	10	US-09-305-736-102	Sequence 102, App
21	51	6.3	1369	11	US-09-818-683-102	Sequence 102, App
22	51	6.3	1369	12	US-10-621-401-102	Sequence 102, App
23	50.4	6.2	854	10	US-09-796-753-49	Sequence 49, Appl
c 24	50.2	6.2	439	9	US-09-917-800A-829	Sequence 829, App
25	49.8	6.1	746	9	US-09-833-503A-5	Sequence 5, Appli
26	49.8	6.1	746	14	US-10-199-881-5	Sequence 5, Appli
27	49.8	6.1	1406	10	US-09-992-600A-81	Sequence 81, Appl
28	49.8	6.1	1406	10	US-09-924-340-81	Sequence 81, Appl
29	49.8	6.1	1406	10	US-09-992-095B-81	Sequence 81, Appl
30	49.8	6.1	1406	10	US-09-999-570-81	Sequence 81, Appl
31	49.8	6.1	1406	14	US-10-000-489-81	Sequence 81, Appl
32	49.8	6.1	1406	14	US-10-000-986-81	Sequence 81, Appl
33	49.8	6.1	1406	14	US-10-154-678-81	Sequence 81, Appl
34	49.8	6.1	1406	14	US-10-001-142-81	Sequence 81, Appl
35	49.8	6.1	1473	10	US-09-796-753-47	Sequence 47, Appl
36	49.8	6.1	1481	14	US-10-198-846-11017	Sequence 11017, A
37	48.8	6.0	423	14	US-10-198-846-8649	Sequence 8649, Ap
c 38	42.2	5.2	559	15	US-10-027-632-197570	Sequence 197570,
39	41.2	5.1	12592	12	US-10-221-613-58	Sequence 58, Appl
40	39.8	4.9	909	14	US-10-156-761-6614	Sequence 6614, Ap
41	39.8	4.9	962	9	US-09-833-503A-3	Sequence 3, Appli
42	39.8	4.9	962	14	US-10-199-881-3	Sequence 3, Appli
43	39.8	4.9	1409	9	US-09-925-301-176	Sequence 176, App
44	39.8	4.9	1422	10	US-09-798-889-24	Sequence 24, Appl
45	39.8	4.9	1422	12	US-10-633-680-24	Sequence 24, Appl

#### ALIGNMENTS

##### RESULT 1

US-09-852-100A-1

; Sequence 1, Application US/09852100A

; Patent No. US20020058267A1

; GENERAL INFORMATION:

; APPLICANT: American Home Products

```

; TITLE OF INVENTION: Beta-amyloid Peptide-Binding Proteins and Polynucleotides
Encoding the
; TITLE OF INVENTION: Same
; FILE REFERENCE: AHP981261p2
; CURRENT APPLICATION NUMBER: US/09/852,100A
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/172,990
; PRIOR FILING DATE: 1998-10-14
; PRIOR APPLICATION NUMBER: US 60/104,104
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: PTC/US99/21621
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: US 09/060,609
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: US 60/064,583
; PRIOR FILING DATE: 1997-04-16
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 810
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(807)
US-09-852-100A-1

```

```

Query Match          100.0%; Score 810; DB 9; Length 810;
Best Local Similarity 100.0%; Pred. No. 1.8e-244;
Matches 810; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 ATGCATATTTTAAAAGGGTCTCCCAATGTGATTCCACGGGCTCACGGGCAGAAGAACACG 60
        |||
Db      1 ATGCATATTTTAAAAGGGTCTCCCAATGTGATTCCACGGGCTCACGGGCAGAAGAACACG 60

Qy     61 CGAAGAGACGGAAC TGGCCTCTATCCTATGCGAGGTCCCTTTAAGAACCTCGCCCTGTTG 120
        |||
Db     61 CGAAGAGACGGAAC TGGCCTCTATCCTATGCGAGGTCCCTTTAAGAACCTCGCCCTGTTG 120

Qy    121 CCCTTCTCCCTCCCGCTCCTGGGCGGAGGCGGAAGCGGAAGTGGCGAGAAAGTGTGCGGTC 180
        |||
Db    121 CCCTTCTCCCTCCCGCTCCTGGGCGGAGGCGGAAGCGGAAGTGGCGAGAAAGTGTGCGGTC 180

Qy    181 TCCAAGATGGCGGCCCGCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGGCCAGA 240
        |||
Db    181 TCCAAGATGGCGGCCCGCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGGCCAGA 240

Qy    241 CTCGTTGGTGTCTGTGGTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGCCACC 300
        |||
Db    241 CTCGTTGGTGTCTGTGGTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGCCACC 300

Qy    301 TCCGCCGGGGGCGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGT 360
        |||
Db    301 TCCGCCGGGGGCGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGT 360

Qy    361 AAAGATCCAAAATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCT 420
        |||

```

Db 361 AAAGATCCAAAAATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCT 420  
 Qy 421 CATGTTTCCTGTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACA 480  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 421 CATGTTTCCTGTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACA 480  
 Qy 481 CATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGC 540  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 481 CATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGC 540  
 Qy 541 TATTCCTACAAAGTGGCAGTCGCATTGTCTCTTTTTCTTGATGGTTGGGAGCAGATCGA 600  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 541 TATTCCTACAAAGTGGCAGTCGCATTGTCTCTTTTTCTTGATGGTTGGGAGCAGATCGA 600  
 Qy 601 TTTTACCTTGGATACCCTGCTTTGGGTTTGTAAAGTTTGCAGTGTAGGGTTTGTGGA 660  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 601 TTTTACCTTGGATACCCTGCTTTGGGTTTGTAAAGTTTGCAGTGTAGGGTTTGTGGA 660  
 Qy 661 ATGGGAGCCTAATTGATTTCAATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGA 720  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 661 ATGGGAGCCTAATTGATTTCAATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGA 720  
 Qy 721 AGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAA 780  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 721 AGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAA 780  
 Qy 781 ACATTTAGAAAAACGCAATTATATCCATAA 810  
 ||||||||||||||||||||||||||||||||  
 Db 781 ACATTTAGAAAAACGCAATTATATCCATAA 810

RESULT 2

US-09-833-503A-1

; Sequence 1, Application US/09833503A

; Patent No. US20020146760A1

; GENERAL INFORMATION:

; APPLICANT: Ozenberger, Bradley A

; APPLICANT: Rajkowski, Eileen M

; APPLICANT: Lo, Ching-Hsiung F

; APPLICANT: American Home Products Corporation

; TITLE OF INVENTION: No. US20020146760A1el G-Protein-Coupled Receptor-Like  
Proteins and

; TITLE OF INVENTION: Polynucleotides Encoded By Them, and Methods of Using

; TITLE OF INVENTION: Same

; FILE REFERENCE: AHP98165-00PCT

; CURRENT APPLICATION NUMBER: US/09/833,503A

; CURRENT FILING DATE: 2000-10-13

; PRIOR APPLICATION NUMBER: 60/104,104

; PRIOR FILING DATE: 1998-10-13

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 810

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-833-503A-1



Query Match 100.0%; Score 810; DB 9; Length 810;  
Best Local Similarity 100.0%; Pred. No. 1.8e-244;  
Matches 810; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGCATATTTTAAAAGGGTCTCCCAATGTGATTCCACGGGCTCACGGGCAGAAGAACACG	60
Db	1	ATGCATATTTTAAAAGGGTCTCCCAATGTGATTCCACGGGCTCACGGGCAGAAGAACACG	60
Qy	61	CGAAGAGACGGAAC TGGCCTCTATCCTATGCGAGGTCCCTTTAAGAACCTCGCCCTGTTG	120
Db	61	CGAAGAGACGGAAC TGGCCTCTATCCTATGCGAGGTCCCTTTAAGAACCTCGCCCTGTTG	120
Qy	121	CCCTTCTCCCTCCCGCTCCTGGGCGGAGGCGGAAGCGGAAGTGGCGAGAAAGTGTCCGGTC	180
Db	121	CCCTTCTCCCTCCCGCTCCTGGGCGGAGGCGGAAGCGGAAGTGGCGAGAAAGTGTCCGGTC	180
Qy	181	TCCAAGATGGCGGCCGCCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGGCCAGA	240
Db	181	TCCAAGATGGCGGCCGCCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGGCCAGA	240
Qy	241	CTCGTTGGTGTCTGTGGTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGCCACC	300
Db	241	CTCGTTGGTGTCTGTGGTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGCCACC	300
Qy	301	TCCGCCGGGGGCGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATATTTGT	360
Db	301	TCCGCCGGGGGCGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATATTTGT	360
Qy	361	AAAGATCCAAAATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACACAGCT	420
Db	361	AAAGATCCAAAATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACACAGCT	420
Qy	421	CATGTTTCCTGTTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACA	480
Db	421	CATGTTTCCTGTTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACA	480
Qy	481	CATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGC	540
Db	481	CATTTTACTGGGAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGC	540
Qy	541	TATTCCTACAAAGTGGCAGTCGCATTGTCTCTTTTTCTTGATGGTTGGGAGCAGATCGA	600
Db	541	TATTCCTACAAAGTGGCAGTCGCATTGTCTCTTTTTCTTGATGGTTGGGAGCAGATCGA	600
Qy	601	TTTTACCTTGGATACCCTGCTTTGGGTTTGTTAAAGTTTGCAGTGTAGGGTTTGTGGA	660
Db	601	TTTTACCTTGGATACCCTGCTTTGGGTTTGTTAAAGTTTGCAGTGTAGGGTTTGTGGA	660
Qy	661	ATTGGGAGCCTAATTGATTTCAATCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGA	720
Db	661	ATTGGGAGCCTAATTGATTTCAATCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGA	720
Qy	721	AGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAA	780
Db	721	AGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAA	780
Qy	781	ACATTTAGAAAAACGCAATTATATCCATAA	810

|||||  
Db 781 ACATTTAGAAAAACGCAATTATATCCATAA 810

RESULT 3

US-10-199-881-1

; Sequence 1, Application US/10199881

; Publication No. US20030096356A1

; GENERAL INFORMATION:

; APPLICANT: Wyeth

; TITLE OF INVENTION: No. US20030096356A1e1 G-Protein-Coupled Receptor-Like  
Proteins and Polynucleotides

; TITLE OF INVENTION: Encoded by Them, and Methods of Using Same"

; FILE REFERENCE: AHP98165C1

; CURRENT APPLICATION NUMBER: US/10/199,881

; CURRENT FILING DATE: 2002-07-18

; PRIOR APPLICATION NUMBER: PCT/ US99/21621

; PRIOR FILING DATE: 1999-10-13

; PRIOR APPLICATION NUMBER: US 90/833,5081

; PRIOR FILING DATE: 2001-12-04

; PRIOR APPLICATION NUMBER: US 60/104,104

; PRIOR FILING DATE: 1998-10-13

; NUMBER OF SEQ ID NOS: 45

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 810

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(810)

; OTHER INFORMATION:

US-10-199-881-1

Query Match 100.0%; Score 810; DB 14; Length 810;

Best Local Similarity 100.0%; Pred. No. 1.8e-244;

Matches 810; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 ATGCATATTTTAAAAGGGTCTCCCAATGTGATTCCACGGGCTCACGGGCAGAAGAACACG 60
          |||
Db      1 ATGCATATTTTAAAAGGGTCTCCCAATGTGATTCCACGGGCTCACGGGCAGAAGAACACG 60

Qy     61 CGAAGAGACGGAACCTGGCCTCTATCCTATGCGAGGTCCCTTTAAGAACCTCGCCCTGTTG 120
          |||
Db     61 CGAAGAGACGGAACCTGGCCTCTATCCTATGCGAGGTCCCTTTAAGAACCTCGCCCTGTTG 120

Qy    121 CCCTTCTCCCTCCGCTCCTGGGCGGAGGCGGAAGCGGAAGTGGCGAGAAAGTGTGCGGTC 180
          |||
Db    121 CCCTTCTCCCTCCGCTCCTGGGCGGAGGCGGAAGCGGAAGTGGCGAGAAAGTGTGCGGTC 180

Qy    181 TCCAAGATGGCGGCCCGCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGGCCAGA 240
          |||
Db    181 TCCAAGATGGCGGCCCGCTGGCCGTCTGGTCCGTCTGCTCCGGAGGCCGTGACGGCCAGA 240

Qy    241 CTCGTTGGTGTCTGTGGTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGCCACC 300
          |||
Db    241 CTCGTTGGTGTCTGTGGTTCGTCTCAGTCACTACAGGACCCTGGGGGGCTGTTGCCACC 300
```

Qy 301 TCCGCCGGGGGCGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATATTTGT 360  
 |||  
 Db 301 TCCGCCGGGGGCGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATATTTGT 360

Qy 361 AAAGATCCAAAAATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACACAGCT 420  
 |||  
 Db 361 AAAGATCCAAAAATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACACAGCT 420

Qy 421 CATGTTTCCTGTTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACA 480  
 |||  
 Db 421 CATGTTTCCTGTTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACA 480

Qy 481 CATTTTACTGGGAACGAAGTTGGTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGC 540  
 |||  
 Db 481 CATTTTACTGGGAACGAAGTTGGTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGC 540

Qy 541 TATTCCTACAAAGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTGGGAGCAGATCGA 600  
 |||  
 Db 541 TATTCCTACAAAGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTGGGAGCAGATCGA 600

Qy 601 TTTTACCTTGGATACCCTGCTTTGGGTTTGTTAAAGTTTTGCACTGTAGGGTTTTGTGGA 660  
 |||  
 Db 601 TTTTACCTTGGATACCCTGCTTTGGGTTTGTTAAAGTTTTGCACTGTAGGGTTTTGTGGA 660

Qy 661 ATTGGGAGCCTAATTGATTTCAATCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGA 720  
 |||  
 Db 661 ATTGGGAGCCTAATTGATTTCAATCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGA 720

Qy 721 AGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAA 780  
 |||  
 Db 721 AGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAA 780

Qy 781 ACATTTAGAAAAACGCAATTATATCCATAA 810  
 |||  
 Db 781 ACATTTAGAAAAACGCAATTATATCCATAA 810

RESULT 4

US-09-922-217-233

; Sequence 233, Application US/09922217

; Patent No. US20020076414A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Lodes, Michael J.

; APPLICANT: Secrist, Heather

; APPLICANT: Benson, Darin R.

; APPLICANT: Meagher, Madeleine Joy

; APPLICANT: Stolk, John A.

; APPLICANT: Wang, Tongtong

; APPLICANT: Jiang, Yuqiu

; APPLICANT: Smith, Carole Lynn

; APPLICANT: King, Gordon E.

; APPLICANT: Wang, Aijun

; APPLICANT: Clapper, Jonathan D.

; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS

; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE

; FILE REFERENCE: 210121.471C13  
; CURRENT APPLICATION NUMBER: US/09/922,217  
; CURRENT FILING DATE: 2001-08-03  
; NUMBER OF SEQ ID NOS: 1124  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 233  
; LENGTH: 508  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-922-217-233

Query Match 61.6%; Score 499; DB 9; Length 508;  
Best Local Similarity 100.0%; Pred. No. 1.8e-146;  
Matches 499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      312 CGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 371
          |||
Db       1 CGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 60

Qy      372 AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACACAGCTCATGTTTCCTG 431
          |||
Db       61 AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACACAGCTCATGTTTCCTG 120

Qy      432 TTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 491
          |||
Db      121 TTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 180

Qy      492 GAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCCTACAA 551
          |||
Db      181 GAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCCTACAA 240

Qy      552 AGTGGCAGTCGCATTGTCTCTTTTTCTTGATGGTTGGGAGCAGATCGATTTTACCTTGG 611
          |||
Db      241 AGTGGCAGTCGCATTGTCTCTTTTTCTTGATGGTTGGGAGCAGATCGATTTTACCTTGG 300

Qy      612 ATACCCTGCTTTGGGTTTGTTAAAGTTTGCAGTGTAGGGTTTGTGGAATTGGGAGCCT 671
          |||
Db      301 ATACCCTGCTTTGGGTTTGTTAAAGTTTGCAGTGTAGGGTTTGTGGAATTGGGAGCCT 360

Qy      672 AATTGATTTTATTCTTATTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 731
          |||
Db      361 AATTGATTTTATTCTTATTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 420

Qy      732 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 791
          |||
Db      421 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 480

Qy      792 AACGCAATTATATCCATAA 810
          |||
Db      481 AACGCAATTATATCCATAA 499
```

RESULT 5

US-09-922-217-245

; Sequence 245, Application US/09922217

; Patent No. US20020076414A1

; GENERAL INFORMATION:

```

; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922,217
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 245
; LENGTH: 508
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-922-217-245

```

```

Query Match          61.6%; Score 499; DB 9; Length 508;
Best Local Similarity 100.0%; Pred. No. 1.8e-146;
Matches 499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      312 CGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 371
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1  CGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 60

Qy      372 AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCTCATGTTTCCTG 431
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCTCATGTTTCCTG 120

Qy      432 TTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 491
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 TTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 180

Qy      492 GAACGAAGTTGGTTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCTTACAA 551
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 GAACGAAGTTGGTTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCTTACAA 240

Qy      552 AGTGGCAGTCGCATTGTCTCTTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG 611
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 AGTGGCAGTCGCATTGTCTCTTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG 300

Qy      612 ATACCCTGCTTTGGGTTTGTAAAGTTTTGCACTGTAGGGTTTTGTGGAATTGGGAGCCT 671
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 ATACCCTGCTTTGGGTTTGTAAAGTTTTGCACTGTAGGGTTTTGTGGAATTGGGAGCCT 360

Qy      672 AATTGATTTTATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 731
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 AATTGATTTTATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 420

```

Qy 732 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 791  
 |||  
 Db 421 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 480  
 Qy 792 AACGCAATTATATCCATAA 810  
 |||  
 Db 481 AACGCAATTATATCCATAA 499

RESULT 6

US-09-833-263-233

; Sequence 233, Application US/09833263

; Patent No. US20020110547A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Aijun

; APPLICANT: Clapper, Jonathan D.

; APPLICANT: Stolk, John A.

; APPLICANT: Meagher, Madeleine J.

; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND

; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE

; FILE REFERENCE: 210121.471C12

; CURRENT APPLICATION NUMBER: US/09/833,263

; CURRENT FILING DATE: 2001-04-10

; NUMBER OF SEQ ID NOS: 1093

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 233

; LENGTH: 508

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-833-263-233

Query Match 61.6%; Score 499; DB 9; Length 508;

Best Local Similarity 100.0%; Pred. No. 1.8e-146;

Matches 499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 312 CGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 371  
 |||  
 Db 1 CGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 60  
 Qy 372 AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCTCATGTTTCCTG 431  
 |||  
 Db 61 AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCTCATGTTTCCTG 120  
 Qy 432 TTTTCCAGCACCCAACATAAATTGTAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 491  
 |||  
 Db 121 TTTTCCAGCACCCAACATAAATTGTAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 180  
 Qy 492 GAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCTTACAA 551  
 |||  
 Db 181 GAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCTTACAA 240  
 Qy 552 AGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG 611  
 |||  
 Db 241 AGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG 300  
 Qy 612 ATACCCTGCTTTGGGTTTGTAAAGTTTTGCACTGTAGGGTTTTGTGGAATTGGGAGCCT 671  
 |||

```

Db      301 ATACCCTGCTTTGGGTTTGTAAAGTTTGGCACTGTAGGGTTTGTGGAATTGGGAGCCT 360
Qy      672 AATTGATTTTCATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 731
      |||
Db      361 AATTGATTTTCATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 420
Qy      732 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 791
      |||
Db      421 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 480
Qy      792 AACGCAATTATATCCATAA 810
      |||
Db      481 AACGCAATTATATCCATAA 499

```

# RESULT 7

US-09-833-263-245

; Sequence 245, Application US/09833263

; Patent No. US20020110547A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Aijun

; APPLICANT: Clapper, Jonathan D.

; APPLICANT: Stolk, John A.

; APPLICANT: Meagher, Madeleine J.

; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND

; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE

; FILE REFERENCE: 210121.471C12

; CURRENT APPLICATION NUMBER: US/09/833,263

; CURRENT FILING DATE: 2001-04-10

; NUMBER OF SEQ ID NOS: 1093

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 245

; LENGTH: 508

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-833-263-245

```

Query Match          61.6%;  Score 499;  DB 9;  Length 508;
Best Local Similarity 100.0%;  Pred. No. 1.8e-146;
Matches 499;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

Qy      312 CGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 371
      |||
Db       1 CGAGGAGTCGCTTAAGTGCGAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 60
Qy      372 AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACACAGCTCATGTTTCCTG 431
      |||
Db       61 AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAACTACACAGCTCATGTTTCCTG 120
Qy      432 TTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 491
      |||
Db      121 TTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 180
Qy      492 GAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCCTACAA 551
      |||
Db      181 GAACGAAGTTGGTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCCTACAA 240

```

Qy 552 AGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG 611  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 241 AGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG 300  
  
 Qy 612 ATACCCTGCTTTGGGTTTGTAAAGTTTGCAGTGTAGGGTTTTGTGGAATTGGGAGCCT 671  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 301 ATACCCTGCTTTGGGTTTGTAAAGTTTGCAGTGTAGGGTTTTGTGGAATTGGGAGCCT 360  
  
 Qy 672 AATTGATTTTCATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 731  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 361 AATTGATTTTCATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 420  
  
 Qy 732 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 791  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 421 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 480  
  
 Qy 792 AACGCAATTATATCCATAA 810  
 ||||||||||||||||  
 Db 481 AACGCAATTATATCCATAA 499

RESULT 8

US-10-025-380-233

; Sequence 233, Application US/10025380

; Publication No. US20020182191A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Lodes, Michael J.

; APPLICANT: Secrist, Heather

; APPLICANT: Benson, Darin R.

; APPLICANT: Meagher, Madeleine Joy

; APPLICANT: Stolk, John A.

; APPLICANT: Wang, Tongtong

; APPLICANT: Jiang, Yuqiu

; APPLICANT: Smith, Carole L.

; APPLICANT: King, Gordon E.

; APPLICANT: Wang, Aijun

; APPLICANT: Clapper, Jonathan D.

; APPLICANT: Skeiky, Yasir A. W.

; APPLICANT: Fanger, Gary R.

; APPLICANT: Vedvick Thomas S.

; APPLICANT: Carter, Darrick

; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS

; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE

; FILE REFERENCE: 210121.471C14

; CURRENT APPLICATION NUMBER: US/10/025,380

; CURRENT FILING DATE: 2001-12-19

; NUMBER OF SEQ ID NOS: 1129

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 233

; LENGTH: 508

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-025-380-233

Query Match 61.6%; Score 499; DB 13; Length 508;

Best Local Similarity 100.0%; Pred. No. 1.8e-146;



Matches 499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      312 CGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 371
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 60

Qy      372 AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCTCATGTTTCCTG 431
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCTCATGTTTCCTG 120

Qy      432 TTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 491
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 TTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 180

Qy      492 GAACGAAGTTGGTTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCCTACAA 551
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 GAACGAAGTTGGTTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCCTACAA 240

Qy      552 AGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG 611
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 AGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG 300

Qy      612 ATACCCTGCTTTGGGTTTGTTAAAGTTTTGCACTGTAGGGTTTTGTGGAATTGGGAGCCT 671
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 ATACCCTGCTTTGGGTTTGTTAAAGTTTTGCACTGTAGGGTTTTGTGGAATTGGGAGCCT 360

Qy      672 AATTGATTTTATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 731
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 AATTGATTTTATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 420

Qy      732 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 791
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 480

Qy      792 AACGCAATTATATCCATAA 810
      ||||||||||||||||
Db      481 AACGCAATTATATCCATAA 499
```

RESULT 9

US-10-025-380-245

; Sequence 245, Application US/10025380

; Publication No. US20020182191A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Lodes, Michael J.

; APPLICANT: Secrist, Heather

; APPLICANT: Benson, Darin R.

; APPLICANT: Meagher, Madeleine Joy

; APPLICANT: Stolk, John A.

; APPLICANT: Wang, Tongtong

; APPLICANT: Jiang, Yuqiu

; APPLICANT: Smith, Carole L.

; APPLICANT: King, Gordon E.

; APPLICANT: Wang, Aijun

; APPLICANT: Clapper, Jonathan D.

; APPLICANT: Skeiky, Yasir A. W.

```
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C14
; CURRENT APPLICATION NUMBER: US/10/025,380
; CURRENT FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 1129
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 245
; LENGTH: 508
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-025-380-245
```

```
Query Match          61.6%; Score 499; DB 13; Length 508;
Best Local Similarity 100.0%; Pred. No. 1.8e-146;
Matches 499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      312 CGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 371
          |||
Db      1   CGAGGAGTCGCTTAAGTGCAGGACCTCAAAGTGGGACAATATATTTGTAAAGATCCAAA 60

Qy      372 AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCTCATGTTTCCTG 431
          |||
Db      61  AATAAATGACGCTACGCAAGAACCAGTTAACTGTACAAACTACACAGCTCATGTTTCCTG 120

Qy      432 TTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 491
          |||
Db      121 TTTTCCAGCACCCAACATAACTTGTAAGGATTCCAGTGGCAATGAAACACATTTTACTGG 180

Qy      492 GAACGAAGTTGGTTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCTTACAA 551
          |||
Db      181 GAACGAAGTTGGTTTTTTTCAAGCCCATATCTTGCCGAAATGTAAATGGCTATTCTTACAA 240

Qy      552 AGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG 611
          |||
Db      241 AGTGGCAGTCGCATTGTCTCTTTTTCTTGGATGGTTGGGAGCAGATCGATTTTACCTTGG 300

Qy      612 ATACCCTGCTTTGGGTTTGTTAAAGTTTTGCACTGTAGGGTTTTGTGGAATTGGGAGCCT 671
          |||
Db      301 ATACCCTGCTTTGGGTTTGTTAAAGTTTTGCACTGTAGGGTTTTGTGGAATTGGGAGCCT 360

Qy      672 AATTGATTTTATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 731
          |||
Db      361 AATTGATTTTATTCTTATTTCAATGCAGATTGTTGGACCTTCAGATGGAAGTAGTTACAT 420

Qy      732 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 791
          |||
Db      421 TATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACATTTAGAAA 480

Qy      792 AACGCAATTATATCCATAA 810
          |||
Db      481 AACGCAATTATATCCATAA 499
```

RESULT 10

US-09-918-995-6918

; Sequence 6918, Application US/09918995

; Publication No. US20030073623A1

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc.

; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED

; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES

; FILE REFERENCE: 20411-756

; CURRENT APPLICATION NUMBER: US/09/918,995

; CURRENT FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: US/09/235,076

; PRIOR FILING DATE: 1999-01-20

; NUMBER OF SEQ ID NOS: 38054

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 6918

; LENGTH: 431

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (1)...(431)

; OTHER INFORMATION: n = A,T,C or G

US-09-918-995-6918

Query Match 41.7%; Score 337.4; DB 10; Length 431;

Best Local Similarity 99.7%; Pred. No. 1.5e-95;

Matches 338; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy      472 AATGAAACACATTTTACTGGGAACGAAGTTGGTTTTTTTCAAGCCCATATCTTGCCGAAAT 531
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  AATGAAACACATTTTACTGGGAACGAAGTTGGTTTTTTTCAAGCCCATATCTTGCCGAAAT 60

Qy      532 GTAAATGGCTATTCCTACAAAGTGGCAGTCGCATTGTCTCTTTTCTTGATGGTTGGGA 591
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db      61  GTAAATGGCTATTCCTACAAAGAGGCAGTCGCATTGTCTCTTTTCTTGATGGTTGGGA 120

Qy      592 GCAGATCGATTTTACCTTGATACCTGCTTTGGGTTTGTTAAAGTTTGCAGTGTAGGG 651
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 GCAGATCGATTTTACCTTGATACCTGCTTTGGGTTTGTTAAAGTTTGCAGTGTAGGG 180

Qy      652 TTTTGTGGAATTGGGAGCCTAATTGATTTATTCTTATTTCAATGCAGATTGTTGGACCT 711
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 TTTTGTGGAATTGGGAGCCTAATTGATTTATTCTTATTTCAATGCAGATTGTTGGACCT 240

Qy      712 TCAGATGGAAGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTATT 771
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 TCAGATGGAAGTAGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTATT 300

Qy      772 ACTAATGAAACATTTAGAAAAACGCAATTATATCCATAA 810
          ||||||||||||||||||||||||||||||||||||
Db      301 ACTAATGAAACATTTAGAAAAACGCAATTATATCCATAA 339

```

RESULT 11

US-10-085-783A-36056

; Sequence 36056, Application US/10085783A

```
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 36056
;   LENGTH: 256
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2)..(2)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (13)..(13)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (30)..(30)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-36056
```

```
Query Match          12.3%; Score 100; DB 12; Length 256;
Best Local Similarity 99.1%; Pred. No. 8.7e-21;
Matches 111; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
```

```
Qy      700 ATTGTTGGACCTTCAG-ATGGAAGTAGTTACATTATAGATTACTATGGAACCAGACTTAC 758
          |||
Db      14  ATTGTTGGACCTTCAGNATGGAAGTAGTTACATTATAGATTACTATGGAACCAGACTTAC 73

Qy      759 AAGACTGAGTATTACTAATGAAACATTTAGAAAAACGCAATTATATCCATAA 810
          |||
Db      74  AAGACTGAGTATTACTAATGAAACATTTAGAAAAACGCAATTATATCCATAA 125
```

---

RESULT 12

```
US-10-242-535A-36056
; Sequence 36056, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
```

```

; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 36056
; LENGTH: 256
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2)..(2)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (13)..(13)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (30)..(30)
; OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-36056

```

```

Query Match          12.3%; Score 100; DB 15; Length 256;
Best Local Similarity 99.1%; Pred. No. 8.7e-21;
Matches 111; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```

```

Qy      700 ATTGTTGGACCTTCAG-ATGGAAGTAGTTACATTATAGATTACTATGGAACCAGACTTAC 758
          |||
Db      14 ATTGTTGGACCTTCAGNATGGAAGTAGTTACATTATAGATTACTATGGAACCAGACTTAC 73

Qy      759 AAGACTGAGTATTACTAATGAAACATTTAGAAAAACGCAATTATATCCATAA 810
          |||
Db      74 AAGACTGAGTATTACTAATGAAACATTTAGAAAAACGCAATTATATCCATAA 125

```

# RESULT 13

US-10-085-783A-48351

; Sequence 48351, Application US/10085783A

; Publication No. US20040037841A1

## ; GENERAL INFORMATION:

; APPLICANT: ChondroGene Inc.

; APPLICANT: Liew, C.C.

; TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis

; FILE REFERENCE: 4231/2002

; CURRENT APPLICATION NUMBER: US/10/085,783A

; CURRENT FILING DATE: 2002-02-28

; PRIOR APPLICATION NUMBER: US 60/305,340

; PRIOR FILING DATE: 2001-07-13

; PRIOR APPLICATION NUMBER: US 60/275,017

; PRIOR FILING DATE: 2001-03-12

```
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 48351
; LENGTH: 411
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-48351
```

Query Match 10.5%; Score 85.4; DB 12; Length 411;  
Best Local Similarity 98.9%; Pred. No. 4.8e-16;  
Matches 86; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy      724 AGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACA 783
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  AGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACA 60

Qy      784 TTTAGAAAAACGCAATTATATCCATAA 810
          |||||||||||| |||||||||
Db      61 TTTAGAAAAACGCAGTTATATCCATAA 87

```

RESULT 14

US-10-242-535A-48351

; Sequence 48351, Application US/10242535A

; Publication No. US20040013663A1

; GENERAL INFORMATION:

; APPLICANT: ChondroGene Inc.

; APPLICANT: Liew, C.C.

: TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis

; FILE REFERENCE: 4231/2005

; CURRENT APPLICATION NUMBER: US/10/242,535A

;  
; CURRENT FILING DATE: 2002-09-12

: PRIOR APPLICATION NUMBER: US 10/085,783

; PRIOR FILING DATE: 2002-02-28

: PRIOR APPLICATION NUMBER: US 60/305,340

; PRIOR FILING DATE: 2001-07-13

: PRIOR APPLICATION NUMBER: US 60/275,017

: PRIOR FILING DATE: 2001-03-12

: PRIOR APPLICATION NUMBER: US 60/271,955

: PRIOR FILING DATE: 2001-02-28

; NUMBER OF SEQ ID NOS: 58994

```

; SOFTWARE: PatentIn version 3.2

```

: SEO ID NO 48351

```

:      .LENGTH: 411

```

TYPE: DNA

: ORGANISM: Human

US-10-242-535A-48351

Query Match 10.5%; Score 85.4; DB 15; Length 411;  
Best Local Similarity 98.9%; Pred. No. 4.8e-16;  
Matches 86; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 724 AGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACA 783  
|||||  
Db 1 AGTTACATTATAGATTACTATGGAACCAGACTTACAAGACTGAGTATTACTAATGAAACA 60

45

US-10-085-783A-16414

US-10-085-783A-16414

Matches 85; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Db

Db

Job time : 351 secs